

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 18:17:14 ; Search time 50.5 Seconds
(without alignments)
1818.375 Million cell updates/sec

Title: US-10-720-192-2
Perfect score: 1726
Sequence: 1 MHQVDNLTTRKGRLLAALAI.....PAGEVAPTTPTPTQTLPA 325

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1726	100.0	325	2 AAW03565	Aaw03565 Mycobacte
2	1726	100.0	325	5 AAW50732	Aam50732 Mycobacte
3	1726	100.0	332	2 AAW32350	Aaw32350 Mycobacte
4	1726	100.0	332	2 AAW32418	Aaw32418 Mycobacte
5	1726	100.0	332	2 AAW43322	Aaw43322 Mycobacte
6	1726	100.0	332	2 AAW81683	Aaw81683 M. tuberc
7	1726	100.0	332	2 AAY38945	Aay38945 M. tuberc
8	1726	100.0	332	2 AAY39083	Aay39083 M. tuberc
9	1726	100.0	332	4 AAU01895	Aau01895 M. tuberc
10	1726	100.0	332	5 AAE29720	Aae29720 Mycobacte
11	1726	100.0	332	5 AAE17584	Aae17584 Mycobacte
12	1720	99.7	325	3 AAY37281	Aay37281 Fibronect
13	1542	89.3	286	2 AAW03566	Aaw03566 Mycobacte
14	1499	86.8	652	2 AAY39082	Aay39082 M. tubercu
15	1499	86.8	652	2 AAY39225	Aay39225 M. tuberc
16	1499	86.8	802	2 AAW64379	Aaw64379 Mycobacte
17	1499	86.8	802	2 AAW81746	Aaw81746 M. tuberc
18	1499	86.8	802	2 AAY32063	Aay32063 Mycobacte
19	1499	86.8	802	2 AAY39033	Aay39033 M. tuberc
20	1499	86.8	802	2 AAY39081	Aay39081 M. tubercu
21	1499	86.8	802	2 AAY39176	Aay39176 M. tuberc
22	1499	86.8	802	2 AAY39224	Aay39224 M. tuberc
23	1499	86.8	802	2 AAY4592	Aau74592 Antigenic
24	980	56.8	287	7 ADB74347	Adb74347 Mycobacte
25	237.5	13.8	268	4 ABUS2942	Abus2942 Human mam

26	235	13.6	256	4 ABUS2945	Abus2945 Human mam
27	233.5	13.5	262	4 ABUS2944	Abus2944 Human mam
28	233	13.5	270	4 ABUS2939	Abus2939 Human mam
29	233	13.5	8991	6 ABUS2938	Abus2938 Human mam
30	228	13.2	544	3 AAG15453	Aag15453 S. pseudom
31	227.5	13.2	652	7 ADB65210	Adb65210 Human pro
32	227	13.2	598	6 ABUS37030	Abus37030 Protein e
33	225.5	13.1	446	4 ABB70063	Abb70063 Drosophil
34	223.5	12.9	652	5 ABB92424	Abb92424 Herbicida
35	222.5	12.9	267	4 ABUS2938	Abus2938 Human mam
36	222	12.9	763	2 AAW31852	Aaw31852 Mycobacte
37	221.5	12.8	260	4 ABUS2932	Abus2932 Human mam
38	221.5	12.8	267	4 ABUS2931	Abus2931 Human mam
39	218.5	12.7	253	4 ABUS2937	Abus2937 Human mam
40	217.5	12.6	817	6 ABR53281	Ab53281 Protein s
41	217	12.6	261	4 ABUS2940	Abus2940 Human mam
42	217	12.6	277	4 ABUS2936	Abus2936 Human mam
43	216.5	12.5	254	4 ABUS2933	Abus2933 Human mam
44	214	12.4	406	4 ABG27250	Abg27250 Novel hum
45	212.5	12.3	258	4 ABUS2941	Abus2941 Human mam

ALIGNMENTS

RESULT 1
AAW03565
ID AAW03565 standard; protein; 325 AA.
XX
AC AAW03565;
XX
DT 22-APR-1997 (first entry)
XX
DE Mycobacterium tuberculosis antigenic determinant protein.

XX Antigenic determinant; Mycobacterium tuberculosis; ion exchange; human;
KW chromatography; gel filtration; reverse phase column chromatography;
KW immunogenic; serum; guinea pig; expression vector; cosmid; antibody;
KW Mycobacterium smegmatis; Mycobacterium bovis BCG; microorganism; vaccine;
KW hybrid; epitope; disease; diphtheria; cholera; toxin.
XX Mycobacterium tuberculosis.

XX Key Location/Qualifiers
FT Peptide 1..39
FT Protein /note= "signal peptide"
FT /note= "mature protein"

XX WO9623885-A1.
XX 08-AUG-1996.
XX 31-JAN-1996; 96WO-FR000166.
XX 01-FEB-1995; 95US-00382184.
XX (INSP) INST PASTEUR.

XX Laqueyzerie A, Marchal G, Pescher P, Romain F;
XX WPI; 1996-371433/37.
XX N-PSDB; AAT39357.

XX Mycobacterium tuberculosis antigen and hybrid proteins comprising it -
XX useful in vaccines against tuberculosis and in immunoassays.
XX Claim 2; Page 49-50; 74pp; French.

XX This is the amino acid sequence of an antigenic determinant protein from
XX Mycobacterium tuberculosis. The mature protein has calculated mol. wt. of
XX 28779 Da but has an experimental mol. wt of 45-47 kD as determined by SDS
XX -PAGE. The difference is thought to be due to the high frequency of Pro

KW skin testing; M.tuberculosis.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO9709429-A2.
 XX
 PD 13-MAR-1997.
 XX
 XX 30-AUG-1996; 96WO-US014675.
 XX
 PR 01-SEP-1995; 95US-00523435.
 PR 22-SEP-1995; 95US-00523436.
 PR 22-MAR-1996; 96US-00620280.
 PR 05-JUN-1996; 96US-00658800.
 PR 12-JUL-1996; 96US-00680573.
 XX
 FA (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
 PI Vedvick TH, Twardzik DR;
 XX
 DR WPI; 1997-192904/17.
 DR N-PSDB; AAT91423.
 XX
 PT New immunogenic polypeptide(s) from soluble M. tuberculosis antigens -
 PT useful for diagnosis of M. tuberculosis infection.
 XX
 PS Example 1; Page 92-94; 190pp; English.
 XX
 CC A new immunogenic polypeptide has been developed comprising an
 CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
 CC variant differing only in conservative substitutions and/or
 CC modifications). The present sequence represents a M.tuberculosis antigen,
 CC DPEP. The immunogenic polypeptide can be used to diagnose M.tuberculosis
 CC infection by forming complexes with specific antibodies in the sample.
 CC fragments of DNA encoding the immunogenic polypeptide can be used as
 CC diagnostic primers or probes and agents that bind to the antigen,
 CC especially monoclonal antibodies or equivalent polyclonal antibodies, are
 CC also used for diagnosis
 XX
 SQ Sequence 332 AA;
 Query Match 100.0%; Score 1726; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 1.3e-104;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MHQVDNLTTRKGRLLAALAIAAMASASLTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
 DB 8 MHQVDNLTTRKGRLLAALAIAAMASASLTVAVPATANADPEPAPPVPTTAASPPSTAAA 67
 QY 61 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
 DB 68 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 127
 QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDDPPFPQPPVANDTRIVLGRDLQKLYASAEA 180
 DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDDPPFPQPPVANDTRIVLGRDLQKLYASAEA 187
 QY 181 TDSKAAARLGSDMGGEFYMPYPTGTRINQETVSLDANGVSGSASYVEVKFSDPSKENGQIWT 240
 DB 188 TDSKAAARLGSDMGGEFYMPYPTGTRINQETVSLDANGVSGSASYVEVKFSDPSKENGQIWT 247
 QY 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVVDKGAALAEISIRPLVAPPAPAPAPAE 300
 DB 248 GVIGSPAANAPDAGPPQRFVWVLTANNPVVDKGAALAEISIRPLVAPPAPAPAPAE 307
 QY 301 APAPAPAGEVAPTPTTPTPTORTLPA 325
 DB 308 APAPAPAGEVAPTPTTPTPTORTLPA 332
 RESULT 4
 AAW32418

ID AAW32418 standard; protein; 332 AA.
 XX
 AC AAW32418;
 XX
 DT 08-JAN-1998 (first entry)
 XX
 DE Mycobacterium tuberculosis antigen DPEP.
 XX
 KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
 KW skin testing; M.tuberculosis.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO9709428-A2.
 XX
 PD 13-MAR-1997.
 XX
 XX 30-AUG-1996; 96WO-US014674.
 XX
 PR 01-SEP-1995; 95US-00523436.
 PR 22-SEP-1995; 95US-00533634.
 PR 22-MAR-1996; 96US-00620874.
 PR 05-JUN-1996; 96US-00659683.
 PR 12-JUL-1996; 96US-00680574.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Skeiky YA, Dillon DC, Campos-Neto A, Houghton R;
 PI Vedvick TH, Twardzik DR;
 XX
 DR WPI; 1997-192903/17.
 DR N-PSDB; AAT91486.
 XX
 PT New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
 PT useful in vaccines for prevention or treatment of tuberculosis, also for
 PT diagnosis.
 XX
 PS Example 1; Page 87-89; 168pp; English.
 XX
 CC A new immunogenic polypeptide has been developed comprising an
 CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
 CC variant differing only in conservative substitutions and/or
 CC modifications). The present sequence represents a M.tuberculosis antigen,
 CC DPEP. The immunogenic protein, and fusion proteins containing one or more
 CC of the proteins or one of the proteins plus ESAT-6, are useful in
 CC vaccines, preferably when formulated with a non-specific adjuvant, to
 CC induce an immune response against M.tuberculosis (for treatment or
 CC prevention)
 XX
 SQ Sequence 332 AA;
 Query Match 100.0%; Score 1726; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 1.3e-104;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MHQVDNLTTRKGRLLAALAIAAMASASLTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
 DB 8 MHQVDNLTTRKGRLLAALAIAAMASASLTVAVPATANADPEPAPPVPTTAASPPSTAAA 67
 QY 61 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
 DB 68 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 127
 QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDDPPFPQPPVANDTRIVLGRDLQKLYASAEA 180
 DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDDPPFPQPPVANDTRIVLGRDLQKLYASAEA 187
 QY 181 TDSKAAARLGSDMGGEFYMPYPTGTRINQETVSLDANGVSGSASYVEVKFSDPSKENGQIWT 240
 DB 188 TDSKAAARLGSDMGGEFYMPYPTGTRINQETVSLDANGVSGSASYVEVKFSDPSKENGQIWT 247
 QY 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVVDKGAALAEISIRPLVAPPAPAPAPAE 300

Db 68 PPAPATVPAPPPAAAANTPNAQPGDPNAPADPNAPPPVIAFNAPQPVRIIDNPVGGF 127
 Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRIVLGRDQKLYASAEA 180
 Db 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRIVLGRDQKLYASAEA 187
 Qy 181 TDSKAAARLGSMDGGEFYMPYFPGTRINQETVSLDANGVSGSASYEVKFSDFSKPENGQIWT 240
 Db 188 TDSKAAARLGSMDGGEFYMPYFPGTRINQETVSLDANGVSGSASYEVKFSDFSKPENGQIWT 247
 Qy 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPPAPAPAPAE 300
 Db 248 GVIGSPAANAPDAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPPAPAPAPAE 307
 Qy 301 APAPAPAGEVAPTPTTPTPQRTPLA 325
 Db 308 APAPAPAGEVAPTPTTPTPQRTPLA 332
 RESULT 7
 AAY38945
 ID AAY38945 standard; protein; 332 AA.
 AC AAY38945;
 XX
 DT 05-NOV-1999 (first entry)
 DE M. tuberculosis recombinant antigen protein DPEP.
 DE Antigen; diagnosis; detection; infection; antibody; immunisation;
 DE vaccine; immunity.
 XX Mycobacterium tuberculosis.
 OS
 XX WO9942118-A2.
 XX
 XX 26-AUG-1999.
 XX
 XX 17-FEB-1999; 99WO-US003265.
 XX
 XX 18-FEB-1998; 98US-00024753.
 PR 05-MAY-1998; 98US-00072596.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
 PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
 DR WPI; 1999-527416/44.
 DR N-PSDB; AAZ19088.
 XX
 PT New polypeptide comprising antigenic portions of M. tuberculosis.
 XX
 XX Example 1; Page 131-132; 323pp; English.
 PS This invention describes novel recombinant antigens and their encoding
 CC nucleic acids derived from Mycobacterium tuberculosis. The novel
 CC polypeptides are useful for detecting M. tuberculosis infection in a
 CC biological sample by detecting antibodies which bind with the
 CC polypeptides, and are useful as vaccines for immunizing against M.
 CC tuberculosis infection. The new detection methods are needed as current
 CC vaccination strategies do not provide 100% immunity
 XX
 SQ Sequence 332 AA;
 Query Match 100.0%; Score 1726; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 1.3e-104;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MEQVDPNLTTRKGRILAALIAAASASILVTAVPATANADPEPAPPPVTTAASPPSTAAA 60
 Db 8 MEQVDPNLTTRKGRILAALIAAASASILVTAVPATANADPEPAPPPVTTAASPPSTAAA 67

Qy 61 PPAPATVPAPPPAAAANTPNAQPGDPNAPADPNAPPPVIAFNAPQPVRIIDNPVGGF 120
 Db 68 PPAPATVPAPPPAAAANTPNAQPGDPNAPADPNAPPPVIAFNAPQPVRIIDNPVGGF 127
 Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRIVLGRDQKLYASAEA 180
 Db 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRIVLGRDQKLYASAEA 187
 Qy 181 TDSKAAARLGSMDGGEFYMPYFPGTRINQETVSLDANGVSGSASYEVKFSDFSKPENGQIWT 240
 Db 188 TDSKAAARLGSMDGGEFYMPYFPGTRINQETVSLDANGVSGSASYEVKFSDFSKPENGQIWT 247
 Qy 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPPAPAPAPAE 300
 Db 248 GVIGSPAANAPDAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPPAPAPAPAE 307
 Qy 301 APAPAPAGEVAPTPTTPTPQRTPLA 325
 Db 308 APAPAPAGEVAPTPTTPTPQRTPLA 332
 RESULT 8
 AAY39083
 ID AAY39083 standard; protein; 332 AA.
 XX
 AC AAY39083;
 XX
 DT 05-NOV-1999 (first entry)
 DE M. tuberculosis antigen DPEP amino acid sequence.
 DE
 DE Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
 XX immunotherapy; diagnosis; immunisation; vaccine; infection;
 XX immune response; skin test.
 XX
 OS Mycobacterium tuberculosis.
 XX
 XX WO9942076-A2.
 XX
 XX 26-AUG-1999.
 XX
 XX 17-FEB-1999; 99WO-US003268.
 XX
 XX 18-FEB-1998; 98US-00025197.
 PR 05-MAY-1998; 98US-00072967.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
 PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
 DR WPI; 1999-527409/44.
 DR N-PSDB; AAZ19300.
 XX
 PT New antigens from Mycobacterium tuberculosis useful in diagnostic skin
 XX tests and protective or therapeutic vaccines or compositions.
 XX
 XX Example 1; Page 93; 299pp; English.
 PS The present invention describes polypeptides comprising an immunogenic
 CC part of a Mycobacterium tuberculosis antigen (Ag). Also described are
 CC vaccines and fusion protein containing M. tuberculosis Ag's. M.
 CC tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other
 CC polypeptides fragments, can be used in pharmaceutical compositions or
 CC vaccines to generate a protective or therapeutic immune response to M.
 CC tuberculosis and as reagents in skin tests for diagnosis of tuberculosis.
 CC Ag can induce proliferation of, or cytokine secretion by, T, B or natural
 CC killer cells and/or macrophages in tuberculosis-immune subjects. AAZ19249
 CC to AAZ19460 and AAY39083 to AAY39225 are used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 332 AA;

Query Match 100.0%; Score 1726; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 1.3e-104;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDPNLTTRKGRLLAALIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
 DB 8 MHQVDPNLTTRKGRLLAALIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 67

QY 61 PPAPATPVAPPPPPAAANTPNAQPGDPNAPPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120
 DB 68 PPAPATPVAPPPPPAAANTPNAQPGDPNAPPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 127

QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFPGQPPVANDTRIVLGRDQKLYASAEA 180
 DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFPGQPPVANDTRIVLGRDQKLYASAEA 187

QY 181 TDSKAAARLGSDMGDFYMPYPCPTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 240
 DB 188 TDSKAAARLGSDMGDFYMPYPCPTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 247

QY 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVDKGAALAESIRPLVAPPAPAPAPAPAP 300
 DB 248 GVIGSPAANAPDAGPPQRFVWVLTANNPVDKGAALAESIRPLVAPPAPAPAPAPAP 307

QY 301 APAPAPAGEVAPTPTTPTPQRTPLA 325
 DB 308 APAPAPAGEVAPTPTTPTPQRTPLA 332

RESULT 9
 AAU01895
 ID AAU01895 standard; protein; 332 AA.
 AC AAU01895;
 DT 29-AUG-2001 (first entry)
 DE M. tuberculosis DPEP antigen.
 DE DPEP; antigen; vaccine; tuberculosis; AIDS;
 KW acquired immunodeficiency disease.
 XX Mycobacterium tuberculosis.
 OS WO200124820-A1.
 PN 12-APR-2001.
 XX 10-OCT-2000; 2000WO-US028095.
 XX 07-OCT-1999; 93US-0158338P.
 PR 07-OCT-1999; 93US-0158425P.
 XX (CORI-) CORIXA CORP.
 PA Skeiky Y, Reed S, Houghton RL, Mcneill PD, Dillon DC, Lodes ML;
 XX WPI; 2001-290576/30.
 DR N-PSDB; AAS03786.
 XX Vaccinating against Mycobacteria infections in mammals using fusion
 PT proteins comprising combinations of heterologous antigens.
 XX Disclosure; Page 160; 168pp; English.
 XX The sequence represents Mycobacterium tuberculosis DPEP, an M.
 CC tuberculosis antigen. Compositions comprising at least 2 heterologous
 CC antigens, as a fusion protein, and vectors expressing the fusion proteins
 CC are used as vaccines to prophylactically immunise mammals (especially
 CC humans) against infection by Mycobacteria. The compositions contain at
 CC least 2 heterologous antigens that increase the serological sensitivity
 CC of individuals infected with tuberculosis, a disease frequently affecting

CC patients with acquired immunodeficiency disease, AIDS
 XX
 SQ Sequence 332 AA;
 Query Match 100.0%; Score 1726; DB 4; Length 332;
 Best Local Similarity 100.0%; Pred. No. 1.3e-104;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDPNLTTRKGRLLAALIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
 DB 8 MHQVDPNLTTRKGRLLAALIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 67

QY 61 PPAPATPVAPPPPPAAANTPNAQPGDPNAPPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120
 DB 68 PPAPATPVAPPPPPAAANTPNAQPGDPNAPPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 127

QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFPGQPPVANDTRIVLGRDQKLYASAEA 180
 DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFPGQPPVANDTRIVLGRDQKLYASAEA 187

QY 181 TDSKAAARLGSDMGDFYMPYPCPTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 240
 DB 188 TDSKAAARLGSDMGDFYMPYPCPTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 247

QY 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVDKGAALAESIRPLVAPPAPAPAPAPAP 300
 DB 248 GVIGSPAANAPDAGPPQRFVWVLTANNPVDKGAALAESIRPLVAPPAPAPAPAPAP 307

QY 301 APAPAPAGEVAPTPTTPTPQRTPLA 325
 DB 308 APAPAPAGEVAPTPTTPTPQRTPLA 332

RESULT 10
 AAE29720
 ID AAE29720 standard; protein; 332 AA.
 XX AAE29720;
 AC AAE29720;
 DT 27-JAN-2003 (first entry)
 DE Mycobacterium tuberculosis DPEP antigenic protein.
 DE Vaccine; immunity; diagnostic agent; gene therapy; DPEP antigen.
 KW Mycobacterium tuberculosis.
 OS WO200272792-A2.
 PN 19-SEP-2002.
 XX 13-MAR-2002; 2002WO-US008223.
 XX 13-MAR-2001; 2001US-0275837P.
 XX (CORI-) CORIXA CORP.
 PA Skeiky Y, Brannon M, Guderian J;
 XX WPI; 2002-759844/82.
 DR N-PSDB; AAD47097.
 XX New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
 PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
 PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium
 CC tuberculosis.
 XX Disclosure; Page 111; 155pp; English.
 PS The invention relates to a recombinant nucleic acid molecule encoding a
 CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
 CC polynucleotide sequence encoding an antigen or an antigenic fragment from
 CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a

CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides, as
CC in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is M.
CC tuberculosis DPEP antigenic protein
XX
SQ Sequence 332 AA;

Query Match 100.0%; Score 1726; DB 5; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.3e-104;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDPNLTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
DB 8 MHQVDPNLTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 67

QY 61 PPAPATPVAPPPAAAANTENAPQDGNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 120
DB 68 PPAPATPVAPPPAAAANTENAPQDGNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 127

QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRI VLGRDLQKLYASAEA 180
DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRI VLGRDLQKLYASAEA 187

QY 181 TDSKAARLGSDMGFEYMPYPTGTRINQETVSLDANGVSGSASYVEVKFSDPSKPNQIWT 240
DB 188 TDSKAARLGSDMGFEYMPYPTGTRINQETVSLDANGVSGSASYVEVKFSDPSKPNQIWT 247

QY 241 GVIGSPAANADAGPPQRFVWVLTANNPVYDKGAALAEIRPLVAPPPAPAPAPAE 300
DB 248 GVIGSPAANADAGPPQRFVWVLTANNPVYDKGAALAEIRPLVAPPPAPAPAPAE 307

QY 301 APAPAGEVAPTPTTPTPQRTLPA 325
DB 308 APAPAGEVAPTPTTPTPQRTLPA 332

RESULT 11
AAE17584
ID AAE17584 standard; protein; 332 AA.
AC AAE17584;
XX
XX
DT 22-APR-2002 (first entry)
DE
DE Mycobacterium species DPEP protein.
KW Fusion protein; antigen; serological sensitivity; immune response;
KW tuberculosis; infection; vaccine; DPEP protein.
XX
XX Mycobacterium sp.
XX
XX WO200198460-A2.
XX
XX 27-DEC-2001.
XX
XX 20-JUN-2001; 2001WO-US019959.
XX
XX 20-JUN-2000; 2000US-00597796.
PR 01-FEB-2001; 2001US-0265737P.
XX
XX (CORI-) CORIXA CORP.
PA
XX
PI Skeiky Y, Reed S, Alderson M;
XX
XX WPI; 2002-147798/19.
DR
DR N-PSDB; AAD28355.
XX
XX Composition comprising MTB39 antigen and MTB32A antigen from

PT Mycobacterium species, useful for eliciting immune response in a subject.
XX
XX Claim 9; Page 127; 136pp; English.
XX
CC The present invention relates to fusion proteins containing at least two
CC Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected with
CC Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is Mycobacterium species
CC DPEP protein
XX
SQ Sequence 332 AA;

Query Match 100.0%; Score 1726; DB 5; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.3e-104;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDPNLTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
DB 8 MHQVDPNLTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 67

QY 61 PPAPATPVAPPPAAAANTENAPQDGNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 120
DB 68 PPAPATPVAPPPAAAANTENAPQDGNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 127

QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRI VLGRDLQKLYASAEA 180
DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRI VLGRDLQKLYASAEA 187

QY 181 TDSKAARLGSDMGFEYMPYPTGTRINQETVSLDANGVSGSASYVEVKFSDPSKPNQIWT 240
DB 188 TDSKAARLGSDMGFEYMPYPTGTRINQETVSLDANGVSGSASYVEVKFSDPSKPNQIWT 247

QY 241 GVIGSPAANADAGPPQRFVWVLTANNPVYDKGAALAEIRPLVAPPPAPAPAPAE 300
DB 248 GVIGSPAANADAGPPQRFVWVLTANNPVYDKGAALAEIRPLVAPPPAPAPAPAE 307

QY 301 APAPAGEVAPTPTTPTPQRTLPA 325
DB 308 APAPAGEVAPTPTTPTPQRTLPA 332

RESULT 12
AAAY97281
ID AAAY97281 standard; protein; 325 AA.
XX
XX AAAY97281;
AC
XX
XX 03-JAN-2001 (first entry)
DT
XX
DE Fibronectin attachment protein FAP-B.
XX
KW Inflammation; inflammatory response; irritant; pathogen; treatment;
KW T helper cell; lymphocyte; cell mediated immunity; skin allergy; hives;
KW allergic rhinitis; conjunctivitis; hay fever; allergic gastroenteritis;
KW asthma; bronchopulmonary aspergillosis; pollutant;
KW respiratory tract infection.

```

OS Mycobacterium bovis.
PN WO200048622-A2.
XX
XX
PD 24-AUG-2000.
XX
XX 22-FEB-2000; 2000WO-US004463.
PF
XX 22-FEB-1999; 99US-0121177P.
PR
XX (IOWA ) UNIV IOWA RES FOUND.
PA
XX Ratliff TL, Kline JN;
PI
XX WPI; 2000-549237/50.
DR
DR N-PSDB; AAA53822.
XX
XX Inhibiting an inflammatory response in a mammal for treating skin
PT allergy, allergic rhinitis, hay fever, or asthma comprises administering
PT a microbial polypeptide.
XX
XX Claim 4; Fig 1; 52pp; English.
PS
XX A new method of treating an inflammatory response comprises administering
XX a microbial polypeptide which is preferably a mycobacterial polypeptide.
CC
CC The method is particularly useful for treating or inhibiting a Th2
CC mediated inflammatory response. The inflammatory response is associated
CC with a disease such as skin allergy, hives, allergic rhinitis,
CC conjunctivitis, hay fever, allergic gastroenteritis, or asthma.
CC Specifically, the asthma is intrinsic, i.e. associated with an irritant
CC (e.g. a pathogen causing a respiratory tract infection in a mammal, or an
CC inhaled pollutant). The asthma may also be extrinsic, which includes
CC allergic asthma, occupational asthma or allergic bronchopulmonary
CC aspergillosis
XX
XX Sequence 325 AA;
Query Match 99.7%; Score 1720; DB 3; Length 325;
Best Local Similarity 99.7%; Pred. No. 3.2e-104;
Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MHQVDNLTTRKGRLLAALAIAMASASLVTVAATANADPEAPPPVTTAASPPSTAAA 60
DB 1 MHQVDNLTTRKGRLLAALAIAMASASLVTVAATANADPEAPPPVTTAASPPSTAAA 60
QY 61 PPAPATPVAPPPAAANTPNAGQDPNAA PPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 120
DB 61 PPAPATPVAPPPAAANTPNAGQDPNAA PPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 120
QY 121 SFALPAGVSDAAHFDYGSALLSKTTGDPFPFGQPPPVANDTRIVLGRDQKLYASAEA 180
DB 121 SFALPAGVSDAAHFDYGSALLSKTTGDPFPFGQPPPVANDTRIVLGRDQKLYASAEA 180
QY 181 TDSKAAARLGSMDGFEYMPYECTRIINQETVSLDANGVSGSASYEVKFSDEPSKENGQIWT 240
DB 181 TDSKAAARLGSMDGFEYMPYECTRIINQETVSLDANGVSGSASYEVKFSDEPSKENGQIWT 240
QY 241 GVIGSPAANADAGPQPPQFWFVVLGTANNPVDKGAALAESIRPLVAPPPAPAPAPAE 300
DB 241 GVIGSPAANADAGPQPPQFWFVVLGTANNPVDKGAALAESIRPLVAPPPAPAPAPAE 300
QY 301 APAPAPAGEVAPPTPTTPTPQRTLPA 325
DB 301 APAPAPAGEVAPPTPTTPTPQRTLPA 325
RESULT 13
AAW03566
ID AAW03566 standard; protein; 286 AA.
XX
XX AAW03566;
XX
XX 22-APR-1997 (first entry)

```

```

XX XX Mycobacterium tuberculosis antigenic determinant mature protein.
DE
XX
XX Antigenic determinant; Mycobacterium tuberculosis; ion exchange; human;
KW chromatography; gel filtration; reverse phase column chromatography;
KW immunogenic; serum; guinea pig; expression vector; cosmid; antibody;
KW Mycobacterium smegmatis; Mycobacterium bovis BCG; microorganism; vaccine;
KW hybrid; epitope; disease; diphtheria; cholera; toxin.
XX
XX Mycobacterium tuberculosis.
OS
XX WO9623885-A1.
PN
XX 08-AUG-1996.
PD
XX 31-JAN-1996; 96WO-FR000166.
PF
XX 01-FEB-1995; 95US-00382184.
PR
XX (INSP ) INST PASTEUR.
PA
XX Laqueyrie A, Marchal G, Pescher P, Romain F;
PI
XX WPI; 1996-371433/37.
DR
DR N-PSDB; AAT39357.
XX
XX Mycobacterium tuberculosis antigen and hybrid proteins comprising it -
PT useful in vaccines against tuberculosis and in immunoassays.
XX
XX Claim 1; Page 49; 74pp; French.
PS
XX This is the amino acid sequence of the mature portion of an antigenic
XX determinant protein from Mycobacterium tuberculosis. The mature protein
XX has calculated mol. wt. of 28779 Da but has an experimental mol. wt. of 45
XX -47 kD as determined by SDS-PAGE. The difference is thought to be due to
XX the high frequency of pro residues (21.7%) in the sequence. The protein
XX was purified from M. tuberculosis strain H37Kv by conventional
XX chromatographic methods e.g. low pressure ion exchange chromatography, Si
XX 300 gel filtration, DEAE ion exchange and reverse phase column
XX chromatography. The proteins were then separated by SDS-PAGE and
XX immunogenic proteins determined with immune serum from humans with TB or
XX guinea pigs infected with M. tuberculosis. A complex of mol. wt. 45-47 kD
XX was identified. An expression library of genomic M. tuberculosis DNA was
XX generated in the cosmid vector pYU18. The library was transformed into
XX M. smegmatis. Clones expressing antigenic determinants were isolated by
XX screening with an antibody against the M. bovis BCG 45/47 kD proteins. 3
XX clones were isolated, all of which contained identical sequences i.e. the
XX sequence presented here. The protein or microorganisms expressing it can
XX be used as vaccines against tuberculosis. Also hybrid proteins comprising
XX this protein and epitopes from other disease causing organisms or
XX proteins, e.g. diphtheria or cholera toxin, can be used as vaccines
XX against their respective diseases
XX
XX Sequence 286 AA;
Query Match 89.3%; Score 1542; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.1e-92;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 DPEAPPPVTTAASPPSTAAAPATPVAPPPAAANTPNAGQDPNAA PPPADPNAPP 99
DB 1 DPEAPPPVTTAASPPSTAAAPATPVAPPPAAANTPNAGQDPNAA PPPADPNAPP 60
QY 100 PVVIAPNAPQVRIIDNPVGGFSFALPAGVSDAAHFDYGSALLSKTTGDPFPFGQPPPV 159
DB 61 PVVIAPNAPQVRIIDNPVGGFSFALPAGVSDAAHFDYGSALLSKTTGDPFPFGQPPPV 120
QY 160 ANDTRIVLGRDQKLYASAEATDSKAAARLGSMDGFEYMPYECTRIINQETVSLDANGVSG 219
DB 121 ANDTRIVLGRDQKLYASAEATDSKAAARLGSMDGFEYMPYECTRIINQETVSLDANGVSG 180
QY 220 SASIYEVKFSDEPSKENGQIWTGVIQSPAANADAGPQPPQFWFVVLGTANNPVDKGAAL 279

```

Db 181 SASIYEVKFSKNGQIWTGVIGSPAANAPDAGPPQRFVWMLGTANNPVDKGAAL 240
Qy 280 AESTRPVAPPPAPAPAPAPAPAPAGVAPTTPTTPTQRTLPA 325
Db 241 AESIRPLVAPPPAPAPAPAPAPAGVAPTTPTTPTQRTLPA 286

RESULT 14
AA39082
ID AA39082 standard; protein; 652 AA.
XX AC AA39082;
XX DT 05-NOV-1999 (first entry)
XX DE M tuberculosis fusion protein Tbf-8.
XX KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
XX KW immunotherapy; diagnosis; immunisation; vaccine; infection;
XX KW immune response; skin test.
XX OS Synthetic.
XX OS Mycobacterium tuberculosis.
XX PN W09942118-A2.
XX PD 26-AUG-1999.
XX PF 17-FEB-1999; 99WO-US003265.
XX PR 18-FEB-1998; 98US-00024753.
XX PR 05-MAY-1998; 98US-00072596.
XX (CORI-) CORIXA CORP.
XX FA Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX XX
XX DR WPI; 1999-527416/44.
XX DR N-PSDB; AA319248.
XX PT New polypeptide comprising antigenic portions of M. tuberculosis.
XX PS Example 10; Page 320-321; 323pp; English.
XX CC This invention describes novel recombinant antigens and their encoding
XX CC nucleic acids derived from Mycobacterium tuberculosis. The novel
XX CC polypeptides are useful for detecting M. tuberculosis infection in a
XX CC biological sample by detecting antibodies which bind with the
XX CC polypeptides, and are useful as vaccines for immunizing against M.
XX CC tuberculosis infection. The new detection methods are needed as current
XX CC vaccination strategies do not provide 100% immunity
XX SQ Sequence 652 AA;

Query Match 86.8%; Score 1499; DB 2; Length 652;
Best Local Similarity 100.0%; Pred. No. 1.5e-89;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 VPTTAASPSSTAAPAPAPATPVAPPPAAANTPNAQPGDPAAPADNAPPPVIAPN 106
Db 374 VPTTAASPSSTAAPAPAPATPVAPPPAAANTPNAQPGDPAAPADNAPPPVIAPN 433
Qy 107 APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFFGQPPPVANDTRIV 166
Db 434 APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFFGQPPPVANDTRIV 493
Qy 167 LGRLDQKLYASAEATDSKAAARLGSMDGFEFYPYPGTRINQETVSLDANGVSGSASYEV 226
Db 494 LGRLDQKLYASAEATDSKAAARLGSMDGFEFYPYPGTRINQETVSLDANGVSGSASYEV 553
Qy 227 KFSDESKNGQIWTGVIGSPAANAPDAGPPQRFVWMLGTANNPVDKGAALAESIRPL 286
Db 554 KFSDESKNGQIWTGVIGSPAANAPDAGPPQRFVWMLGTANNPVDKGAALAESIRPL 613

Qy 287 VAPPPAPAPAPAPAPAPAPAGVAPTTPTTPTQRTLPA 325
Db 614 VAPPPAPAPAPAPAPAPAPAGVAPTTPTTPTQRTLPA 652

RESULT 15
AA39225
ID AA39225 standard; protein; 652 AA.
XX AC AA39225;
XX DT 05-NOV-1999 (first entry)
XX DE M. tuberculosis fusion protein Tbf-8 amino acid sequence.
XX KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
XX KW immunotherapy; diagnosis; immunisation; vaccine; infection;
XX KW immune response; skin test.
XX OS Synthetic.
XX OS Mycobacterium tuberculosis.
XX PN W09942076-A2.
XX PD 26-AUG-1999.
XX PF 17-FEB-1999; 99WO-US003268.
XX PR 18-FEB-1998; 98US-00025197.
XX PR 05-MAY-1998; 98US-00072967.
XX (CORI-) CORIXA CORP.
XX FA Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX XX
XX DR WPI; 1999-527409/44.
XX DR N-PSDB; AA219460.
XX PT New antigens from Mycobacterium tuberculosis useful in diagnostic skin
XX PS tests and protective or therapeutic vaccines or compositions.
XX SQ Claim 37; Page 274-276; 299pp; English.

The present invention describes polypeptides comprising an immunogenic part of a Mycobacterium tuberculosis antigen (Ag). Also described are vaccines and fusion protein containing M. tuberculosis Ag's. M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other polypeptides fragments, can be used in pharmaceutical compositions or vaccines to generate a protective or therapeutic immune response to M. tuberculosis and as reagents in skin tests for diagnosis of tuberculosis. Ag can induce proliferation of, or cytokine secretion by, T, B or natural killer cells and/or macrophages in tuberculosis-immune subjects. AA219249 to AA219460 and AA39083 to AA39225 are used in the exemplification of the present invention

Query Match 86.8%; Score 1499; DB 2; Length 652;
Best Local Similarity 100.0%; Pred. No. 1.5e-89;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 VPTTAASPSSTAAPAPAPATPVAPPPAAANTPNAQPGDPAAPADNAPPPVIAPN 106
Db 374 VPTTAASPSSTAAPAPAPATPVAPPPAAANTPNAQPGDPAAPADNAPPPVIAPN 433
Qy 107 APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFFGQPPPVANDTRIV 166
Db 434 APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFFGQPPPVANDTRIV 493
Qy 167 LGRLDQKLYASAEATDSKAAARLGSMDGFEFYPYPGTRINQETVSLDANGVSGSASYEV 226
Db 494 LGRLDQKLYASAEATDSKAAARLGSMDGFEFYPYPGTRINQETVSLDANGVSGSASYEV 553

Qy	227	KFSDESKPNCQIWTGVI	GSPAANAPDAGPPORWFV	VLGTANNPVVDKGAAYALAESIRPL	286
Db	554	KFSDESKPNCQIWTGVI	GSPAANAPDAGPPORWFV	VLGTANNPVVDKGAAYALAESIRPL	613
Qy	287	VAPPPAPAPAPAPAPAP	AGEVAPTPTTTPORTLPA	325	
Db	614	VAPPPAPAPAPAPAPAP	AGEVAPTPTTTPORTLPA	652	

Search completed: July 7, 2004, 18:21:03
Job time : 53.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 18:19:15 ; Search time 22.5 Seconds
(without alignments)
1389.433 Million cell updates/sec

Title: US-10-720-192-2
Perfect score: 1726
Sequence: 1 MHQVDNLTTRKGRGLAALAI.....PAGEVAPTPTPTPQRTLPA 325

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pirl1:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1726	100.0	325	2 D70666	probable modD prot
2	980	56.8	287	2 B87166	hypothetical prote
3	241	14.0	1585	2 T31611	hypothetical prote
4	237.5	13.8	1188	2 S49915	extensin-like prot
5	228.5	13.2	464	2 S22697	extensin - Volvox
6	226.5	13.1	222	2 H96711	hypothetical prote
7	224.5	13.0	3534	2 T42567	tegument protein 2
8	219.5	12.7	85	2 T44744	antigen 43L [impor
9	217.5	12.6	817	2 S51342	verprolin - yeast
10	217.5	12.6	839	2 T75518	hypothetical prote
11	217	12.6	801	2 T29018	hypothetical prote
12	216.5	12.5	847	2 F96531	hypothetical prote
13	213.5	12.4	760	2 T06291	extensin homolog T
14	213.5	12.4	3164	1 WMBEHS	UL36 protein - hum
15	208.5	12.1	214	2 T10737	extensin-like cell
16	207	12.0	416	1 SXXLAG	dermal gland prote
17	206.5	12.0	1611	2 T38236	hypothetical prote
18	204.5	11.8	539	2 T28770	hypothetical prote
19	203	11.8	1201	2 G86441	unknown protein [i
20	203	11.8	3421	1 W29EB6	367K tegument prot
21	202.5	11.7	283	2 S13383	hydroxyproline-ric
22	201.5	11.7	395	2 H75457	hypothetical prote
23	200.5	11.6	214	2 T09854	proline-rich cell
24	200.5	11.6	760	2 F86387	probable Pto kinas
25	199	11.5	708	2 D96711	hypothetical prote
26	199	11.5	839	2 T04859	extensin homolog F
27	197.5	11.4	865	2 A47282	calcium-binding pr
28	196.5	11.4	418	2 T15142	hypothetical prote
29	196.5	11.4	666	2 B70803	hypothetical prote

30	194	11.2	191	2 F84522	probable proline-r
31	193.5	11.2	1151	2 T18535	high molecular mas
32	193	11.2	620	2 S06733	hydroxyproline-ric
33	192.5	11.2	731	2 T04455	hypothetical prote
34	192.5	11.2	744	2 E86255	hypothetical prote
35	192.5	11.2	873	2 A47283	calphostin fruit
36	191.5	11.1	280	2 T11671	extensin-like prot
37	191	11.1	576	2 T36729	probable serine/th
38	190	11.1	550	2 G70597	probable proteinas
39	189.5	11.0	267	2 S08314	cell wall glycopro
40	189	11.0	240	2 B24264	proline-rich prote
41	188	11.0	1794	2 T38459	hypothetical diver
42	188	10.9	544	2 T17547	proline-rich prote
43	188	10.9	699	2 C43674	US4 protein - huma
44	188	10.9	1268	2 T31420	C-terminal domain-
45	187	10.8	350	2 S22456	hydroxyproline-ric

ALIGNMENTS

RESULT 1

D70666
probable modD protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C/Accession: D70666
R/Cole, S.T.; Brogi, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537(544), 1998
A/Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the Biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70360; MUID:98295987; PMID:9634230
A/Accession: D70666
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-325 <COL>
A/cross-references: GB:Z83859; GB:AL123456; NID:G3261678; PIDN:CAB06127.1; PID:e290722, I
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: modD

Query Match	100.0%	Score 1726;	DB 2;	Length 325;
Best Local Similarity	100.0%	Pred. No. 2.8e-87;		
Matches 325;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MHQVDNLTTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPVPVPTTASPSTAAA	60	
Db	1	MHQVDNLTTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPVPVPTTASPSTAAA	60	
Qy	61	PPAPATFVAPPPPPAAANTPNAOPGDPNAPPPVIAPNAPOPVIDNPVGGF	120	
Db	61	PPAPATFVAPPPPPAAANTPNAOPGDPNAPPPVIAPNAPOPVIDNPVGGF	120	
Qy	121	SFALPAGWESDAAHFDYGSALLSKTTGPPPGPPGVANDTRIVLGLDQKLVASABA	180	
Db	121	SFALPAGWESDAAHFDYGSALLSKTTGPPPGPPGVANDTRIVLGLDQKLVASABA	180	
Qy	181	TDSKAARLGSDMGREYMPYCGTRINQETVSDANGVSGASVYEVKFSDDPSKNGQIWT	240	
Db	181	TDSKAARLGSDMGREYMPYCGTRINQETVSDANGVSGASVYEVKFSDDPSKNGQIWT	240	
Qy	241	GVIGSPAANAPDAGPPQRFVFWLGTANNPVDKGAALAESIRPLVAPPPAPAPAPAP	300	
Db	241	GVIGSPAANAPDAGPPQRFVFWLGTANNPVDKGAALAESIRPLVAPPPAPAPAPAP	300	
Qy	301	APAPAPAGEVAPTPTPTPTQRTLPA	325	
Db	301	APAPAPAGEVAPTPTPTPTQRTLPA	325	

RESULT 2

D _b	1293	A P A P A S S G Y S G G S G S A A C G G G S S G Y S G S A A P P P P P P P P P A P - A P A - S	1350
Q _y	113	I N F V G G S F A L P A Q W E S D A H F Y G S A L K T T G D P P F Q C Q P P P V A N D R I V L G R L D Q	172
D _b	1351	S G Y S G G S G S A A G G G S S G Y T G S A - A P P P P P P P P P P P A P -	1395
Q _y	173	K U Y A B A T D S K A A R L G S D M G E F Y M P Y P G T R I N O E T V S L D A N G V S G S A Y Y E V K F S D P S	232
D _b	1396	- - - A P A P A S S G Y S G - G S S G S - - - A A G G G G S G Y S G S A A P P	1434
Q _y	233	K P N G I W T G I V G S P A A N A D A G P P O R F V W L G T A N N P V D K A A X A - - - L A E S I R P L V A	288
D _b	1435	P P - - - - - P P A P A P A P S G G Y S G - - G S G G S A A G G G S S G G Y T G C S A A P - - -	1478
Q _y	289	P P P A P A P A E P A P A P A - - - - - G E V A P T P T T P T Q	320
D _b	1479	P P P P P P P P P A P A P A P A P A P S G G Y S G S G S A A G G G G S S G Y T G S A A P P P P P P P P	1538

```

QY      321 RUPA 325
      ||
Db      1539 PPPPA 1543

RESULT 4
S49915
  extensin-like protein - maize
C:Species: Zea mays (maize)
C:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 29-Oct-1999
C:Accession: S49915
R:Rubinstein, A.L.; Broadwater, A.H.; Lowrey, K.; Bedinger, P.A.
  submitted to the EMBL Data Library, June 1994
A:Description: Pex genes: pollen-specific genes with extensin-like domains.
A:Reference number: S49915
A:Accession: S49915
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1188 <RUB>
A:Cross-references: EMBL:234465; NID:G600117; PIDN:CAA84230.1; PID:G600118

Query Match      13.8%; Score 237.5; DB 2; Length 1188;
Best Local Similarity 28.9%; Pred. No. 8.2e-06;
Matches 87; Conservative 30; Mismatches 123; Indels 61; Gaps 13;

```

[illegible]

Qy 324 P 324
Db 818 P 818

RESULT 5
S22697
extensin - Volvox carteri (fragment)
C.Species: Volvox carteri

C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 21-Jul-2000
 C;Accession: S22697; S21006
 R;Ertl, H.; Hallmann, A.; Wenzl, S.; Sumper, M.
 EMBO J. 11, 2055-2062, 1992
 A;Title: A novel extension that may organize extracellular matrix biogenesis in *Volvox carterii*
 A;Reference number: S22697; PMID:92289669; PMID:1600938
 A;Accession: S22697
 A;Molecule type: mRNA
 A;Residues: 1-464 <HL>
 A;Cross-references: EMBL:X65165; NID:g21991; PIDN:CAA46283.1; PID:g21992
 C;Keywords: Glycoprotein

Query Match 13.2%; Score 228.5; DB 2; Length 464;
 Best Local Similarity 22.3%; Pred. No. 1e-05;
 Matches 70; Conservative 18; Mismatches 69; Indels 157; Gaps 9;

QY 34 PATANADPEA-----PPVPTTAASPPSTAAAPPAPATVA--PPP 72
 DB 235 PARVSSPPATRRPPRRITSPVLTASPLEKT--SPPPRRVPPSPPPVAGSPPPP 292
 QY 73 PAANTNAOGDPAADPADP---NAPPPVIAADNAPQVRIIDNPVGGFALPAGW 128
 DB 293 PPRVSPSPPPQVSPSPPPPPRPSPPPPRSPSPSPPP----- 334
 QY 129 VESDAAHFDYGSALLSKTTGDPFPQPPVANDTRIVLGRDQKLYASAEATDSKAAAR 188
 DB 335 -----SPPPSPPPPPRSPS----- 349
 QY 189 LGSDMGFYMPYPCTRINQETVSLDANGVSGSASYEVKPSDKNGQIWTGVIGSPAA 248
 DB 350 -----PPPPR-----SSPSPPPPV-----SPPP 368
 QY 249 NADGAPQPFQFVWVWLTGANNPVDKGAALAEISIRPLVAPPAPAPAPAPAPAG 308
 DB 369 PPRASPPP-----PPASPPPPPPPPPPPPPPPPPPPPA 402
 QY 309 EVAPTPPTTPTPQRT 322
 DB 403 TAAANPSPAPSR 416

RESULT 6
 H96711
 hypothetical protein Flak14.17 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: H96711
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A;Reference number: A86141; PMID:21016719; PMID:11130712
 A;Accession: H96711
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-222 <STO>
 A;Cross-references: GB:AB005173; NID:g6524187; PIDN:AAF15072.1; GSPDB:GN00141
 C;Genetics:
 A;Gene: Flak14.17
 A;Map position: 1

Query Match 13.1%; Score 226.5; DB 2; Length 222;
 Best Local Similarity 23.1%; Pred. No. 6.2e-06;
 Matches 75; Conservative 14; Mismatches 75; Indels 161; Gaps 9;

QY 22 AVASASLVTVAVATANADPPAPVPTT-----AASPPSTAAAPPAPATPVAPPAAA 75

DB 28 AASPVTSTTTAPPTTAAPPTTAAAPPTTTTPPVSAQPPASPVTTPPPAVTPTSPAPKV 87
 QY 76 AN--TPNAQPGDPAAPADPNAPPVIAAPNAPQVRIIDNPVGGFALPAGWVESDA 133
 DB 88 ADVISPATPPQPPSPASAPTTSPPVSPPPAP----- 122
 QY 134 AHFDYGSALLSKTTGDPFPQPPVANDTRIVLGRDQKLYASAEATDSKAAARLGSDM 193
 DB 123 -----TSPPPTPASPP-----APASPPPPA----- 144
 QY 194 GFYMPYPCTRINQETVSLDANGVSGSASYEVKPSDKNGQIWTGVIGSPAANAPDA 253
 DB 145 -----SPPAP-----VSPPPVQAPSP 161
 QY 254 GPPQRFVWVWLTGANNPVDKGAALAEISIRPLVAPPAPAPAPAE-----P 300
 DB 162 -----ISLPAPAPAPTKHKRKHKEHHAP 188

RESULT 7
 T42567
 tegument protein 24 - equine herpesvirus 4 (strain NS80567)
 C;Species: equine herpesvirus 4
 A;Variety: strain NS80567
 C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
 R;Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
 J. Gen. Virol. 79, 1197-1203, 1998
 A;Title: The DNA sequence of equine herpesvirus-4.
 A;Reference number: Z22173; PMID:98264497; PMID:9603335
 A;Accession: T42567
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-3534 <TEL>
 A;Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AAC59539.1; PID:g2605967
 C;Experimental source: strain NS80567
 C;Genetics:
 A;Note: 24
 C;Superfamily: varicella-zoster virus gene 22 protein

Query Match 13.0%; Score 224.5; DB 2; Length 3534;
 Best Local Similarity 27.8%; Pred. No. 0.00012;
 Matches 91; Conservative 18; Mismatches 93; Indels 125; Gaps 13;

QY 5 DPNLRRKGRGLAALAAAMASASLVTVAVATANADPPAPVPTTAAAP--PSTAAAPP 62
 DB 2710 DPN-----EALLTAPSKPAAAPAPSKPAAAPAPSKPAAAPAPSKPAAAPAPSK 2765
 QY 63 APATPVAPPPPAAANTPNAPQDPNAPPPADPNAPPVIAAPNAPQVRIIDNPVGGFSF 122
 DB 2766 KPAAPAPSKPAAAPAPSK-----KPAAPAPSKPAAAPAPSKPAAAPAPSK----- 2811
 QY 123 ALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPVANDTRIVLGRDQKLYASAEATD 182
 DB 2812 --PAA-----APAPSKP-----AAAPAPSK 2828
 QY 183 SKAAARLGSDMGFYMPYPCTRINQETVSLDANGVSGSASYEVKPSDKNGQIWTGV 242
 DB 2829 KPAAP-----PAPSKP-----PAPSKP----- 2839
 QY 243 IGSPAANADAGP--PQRFVWVWLTGANNPVDKGAALAEISIRPLVAP---PPAPAPAP 298
 DB 2840 AAAPAPSKPAAAPAPSK-----PAAAPAPSKPAAAPAPSKPAAAPAPSKPAAAPAPSK 2891
 QY 299 EPAPAPAPAGEVAPTPTTPTPQRTLPA 325
 DB 2892 KPAAPAPSKPAAAPAPSKPQNTILVA 2917

RESULT

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 18:17:39 ; Search time 10.5 Seconds
(without alignments)
1611.695 Million cell updates/sec

Title: US-10-720-192-2

Perfect score: 1726

Sequence:

1 MHQVDPNLTTRKGRKLAALAI.....PAGEVAPTPTPTPTQTLPA 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1726	100.0	325	1	APA_MYCTU
2	1720	99.7	325	1	APA_MYCTU
3	985	57.1	381	1	APA_MYCAV
4	980	56.8	287	1	APA_MYCLE
5	243	14.1	555	1	GPI_CHLRE
6	217.5	12.6	817	1	VRP1_YEAST
7	213.5	12.4	3164	1	TEGU_HSV11
8	210	12.2	1083	1	T2D3_HUMAN
9	207	12.0	439	1	XP2_XENLA
10	205.5	11.9	802	1	ENAH_MOUSE
11	203	11.8	3421	1	TEGU_HSVB
12	197.5	11.4	865	1	CPN_DROME
13	197	11.4	283	1	EXTN_SORBI
14	193	11.2	620	1	EXTN_TOBAC
15	192.5	11.2	1198	1	HCN4_RAT
16	191	11.1	1386	1	ZAP3_MOUSE
17	189.5	11.0	267	1	EXTN_MAIZE
18	189	11.0	1794	1	YDC9_SCHPO
19	189	11.0	2167	1	SHK1_RAT
20	188.5	10.9	2715	1	MLL4_HUMAN
21	188	10.9	699	1	VGLG_HSV2H
22	186.5	10.8	5179	1	MUC2_HUMAN
23	186	10.8	331	1	PRP1_HUMAN
24	186	10.8	1033	1	IF2_STRCO
25	185.5	10.7	1300	1	SAL3_HUMAN
26	184.5	10.7	639	1	ZIC5_HUMAN
27	184.5	10.7	1046	1	IF2_STRAW
28	184	10.7	1271	1	IF2_STRAW
29	183	10.6	347	1	CSP_PLABA
30	182	10.5	261	1	PRP2_MOUSE
31	180.5	10.5	534	1	APG_ARATH
32	180.5	10.5	2716	1	OSA_DROME
33	179.5	10.4	2142	1	BAT2_HUMAN

34	179	10.4	339	1	CSP_PLABE
35	178	10.3	245	1	CM32_HUMAN
36	177.5	10.3	1575	1	SYJ1_HUMAN
37	177	10.3	1248	1	DIA1_HUMAN
38	177	10.3	2161	1	SHK1_HUMAN
39	176.5	10.2	290	1	TRX2_MOUSE
40	176.5	10.2	1132	1	BAT3_HUMAN
41	175	10.1	296	1	PMP3_MOUSE
42	174.5	10.1	502	1	WASP_HUMAN
43	174.5	10.1	639	1	SF01_HUMAN
44	174.5	10.1	1324	1	IRS2_HUMAN
45	174.5	10.1	2517	1	NCR2_HUMAN

ALIGNMENTS

RESULT 1

APA_MYCTU

ID APA_MYCTU STANDARD; PRT; 325 AA.

AC Q50906; O08062; 39, Created

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Alanine and proline-rich secreted protein (a precursor (45/47 kDa antigen) (Fibronectin attachment protein) (Immunogenic protein MPT32) (Antigen MPT-32) (45-kDa glycoprotein) (FAP-B).

GN APA OR MODD OR RV1860 OR M1908 OR MTCY359.13.

OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID:1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37Rv;

RX MEDLINE=96009758; PubMed=7558311;

RA Laquerriere A., Milltzer P., Romain F., Eiglmeier K., Cole S.,

RA Marchel G.;

RT Mycobacterium tuberculosis 45/47-kilodalton secreted antigen

RL Infect. Immun. 63:4003-4010(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=H37Rv;

RA Laquerriere A.;

RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=H37Rv;

RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Seesch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmeier K., Gas S., Barry C.B. III, Tekala F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Sulten J., Seeger K., Skelton S., Squares S., Squares R.,

RA Sulten J., Taylor K., Whitehead S., Barrell B.G.;

RT Deciphering the biology of Mycobacterium tuberculosis from the

RL complete genome sequence.

RL Nature 393:537-544(1998).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=CDC 1551; Oshkosh;

RX MEDLINE=22206494; PubMed=12218036;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey B.,

RA Kolarov J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,

RA Delcher A., Uterback T., Weidman J.C., Khouri H., Gill J., Mikula A.,

RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;

RT Whole-genome comparison of Mycobacterium tuberculosis clinical and

RT laboratory strains."

Handwritten signature

RL J. Bacteriol. 184:5479-5490(2002).
RN [5]
RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=9109989; PubMed=1898899;
RA Nagai S., Wiker H.G., Harboe M., Kinomoto M.;
RT "Isolation and partial characterization of major protein antigens in
RL the culture fluid of Mycobacterium tuberculosis";
RN Infect. Immun. 59:372-382(1991).
RX [6]
RP PARTIAL SEQUENCE, AND GLYCOSYLATION.
RX MEDLINE=95347792; PubMed=7622204;
RA Dobos K.M., Swiderek K., Khoo K.-H., Brennan P.J., Belisle J.T.;
RT "Evidence for glycosylation sites on the 45-kilodalton glycoprotein of
RL Mycobacterium tuberculosis";
RN Infect. Immun. 63:2846-2853(1995).
RX [7]
RP PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=96196153; PubMed=8626314;
RA Dobos K.M., Khoo K.-H., Swiderek K.M., Brennan P.J., Belisle J.T.;
RT "Definition of the full extent of glycosylation of the 45-kilodalton
RL glycoprotein of Mycobacterium tuberculosis";
RN J. Bacteriol. 178:2498-2506(1996).
RX [8]
RP SEQUENCE OF 40-49, AND CHARACTERIZATION OF CARBOHYDRATE-LINKAGE SITES
BY MASS SPECTROMETRY.
RX STRAIN=H37RV;
RX MEDLINE=20011399; PubMed=10542234;
RA Horn C., Namane A., Pescher P., Riviere M., Romain F., Puzo G.,
Barzu O., Marchal G.;
RT "Decreased capacity of recombinant 45/47-kDa molecules (Apa) of
RL Mycobacterium tuberculosis to stimulate T lymphocyte responses
related to changes in their mannoseylation pattern";
RN J. Biol. Chem. 274:32023-32030(1999).
RX [9]
RP SEQUENCE OF 40-57, AND CHARACTERIZATION OF CARBOHYDRATE-LINKAGE SITES
BY MASS SPECTROMETRY.
RX STRAIN=H37RV;
RX MEDLINE=20002527; PubMed=10531201;
RA Romain F., Horn C., Pescher P., Namane A., Riviere M., Puzo G.,
Barzu O., Marchal G.;
RT "Deglycosylation of the 45/47-kilodalton antigen complex of
RL Mycobacterium tuberculosis decreases its capacity to elicit in vivo
or in vitro cellular immune responses";
RN Infect. Immun. 67:5567-5572(1999).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MASS SPECTROMETRY: MW=28782; MW_ERR=1.6; METHOD=Electrospray;
CC RANGE=40-325.
CC -!- BIOTECHNOLOGY: Major immunodominant antigen that has potential as
CC a vaccine against tuberculosis. APA-ELISA could be used in
CC diagnosis.
CC -!- MISCELLANEOUS: Changes in the mannoseylation pattern of this
CC protein affect its ability to stimulate T-lymphocyte response.
CC -!- CAUTION: Was originally thought to be involved in molybdenum
CC transport.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X80268; CA856555.1; -
DR EMBL; X99458; CA867645.1; -
DR EMBL; Z83859; CAB06127.1; -
DR EMBL; AB007048; AAK46179.1; -
DR PIR; D70666; D70666.
DR TIGR; MT1908; -
DR Tuberculist; Rv1860; -
DR Antigen; Glycoprotein; Repeat; Signal; Complete proteome.
FT SIGNAL 1 39
FT CHAIN 40 325 ALANINE AND PROLINE-RICH SECRETED PROTEIN

PT DOMAIN 85 107 APA.
FT 3 X 4 AA APPROXIMATE REPEATS OF [DA]-P-
FT N-A.
FT REPEAT 85 88 1.
FT REPEAT 94 97 2.
FT REPEAT 104 107 3.
FT CARBOHYD 49 49 O-LINKED (MAN. . .).
FT CARBOHYD 57 57 O-LINKED (MAN. . .).
FT CARBOHYD 66 66 O-LINKED (MAN. . .).
FT CARBOHYD 316 316 O-LINKED (MAN. . .).
FT CONFLICT 136 136 F -> L (IN REF. 4).
SQ SEQUENCE 325 AA; 32720 MW; 59B5D0455A997BED CRC64;
Query Match 100.0%; Score 1726; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.5e-80;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEQVDPNLTTRKGRLLAALIAAMASASLVTVAVPATANADPEAPVPVPTTAASPPSTAAA 60
DB 1 MEQVDPNLTTRKGRLLAALIAAMASASLVTVAVPATANADPEAPVPVPTTAASPPSTAAA 60
QY 61 PPAPATPVAPPPAAANTPNAQPGDPNAPPPADPNAPPPVIAAPNAPQVRIDNPVGGF 120
DB 61 PPAPATPVAPPPAAANTPNAQPGDPNAPPPADPNAPPPVIAAPNAPQVRIDNPVGGF 120
QY 121 SPALPAGWVESDAAHFDYGSALLSKTTGDPGGPPGPPVANDTRIVLGRLOKLYASAEA 180
DB 121 SPALPAGWVESDAAHFDYGSALLSKTTGDPGGPPGPPVANDTRIVLGRLOKLYASAEA 180
QY 181 TTSKAAARLGSDMGSEFYMPYPTGTRINQETVSLDANGVSGSYYEVKTSKPNQIWT 240
DB 181 TTSKAAARLGSDMGSEFYMPYPTGTRINQETVSLDANGVSGSYYEVKTSKPNQIWT 240
QY 241 GVIGSFAANADAGPPQRFVVMVLTANNPVDKGAALAESIRPLVAPPAPAPAPAP 300
DB 241 GVIGSFAANADAGPPQRFVVMVLTANNPVDKGAALAESIRPLVAPPAPAPAPAP 300
QY 301 APAPAGEVAPT 325
DB 301 APAPAGEVAPT 325
RESULT 2
APA_MYCBO STANDARD; PRT; 325 AA.
ID APA_MYCBO
AC G30620;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Alanine and proline-rich secreted protein apa precursor (Fibronectin
DE attachment protein) (45/47 kDa antigen) (FAP-B).
GN APA OR ME1891.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCG;
RA Zhao W., Shorey J.S., Bong-Mastek M., Brown E.J., Ratliff T.L.;
RT "Identification, sequence and characterization of the M. bovis SCG
RT fibronectin attachment protein";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeyer K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).

RP SEQUENCE FROM N.A.
RX MEDLINE=94086110; PubMed=8262636;


```
RP SEQUENCE FROM N.A.
RX MEDLINE=8274327; PubMed=2839594;
RA McGeoch D.J., Dairymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1."
RL J. Gen. Virol. 69:1531-1574 (1998).
CC -!- FUNCTION: Tegument protein.
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
CC EH-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; X14112; CAA32311.1; -.
DR PIR; I30085; WMBE6.
DR InterPro; IPR006928; Herpes teg N.
DR InterPro; IPR005210; Herpes UL36.
DR Pfam; PF04843; Herpes teg N; 1.
DR Pfam; PF03586; Herpes_UL36; 1.
KW Repeat.
FT DOMAIN
SQ SEQUENCE 2911 2980 35 X 2 AA TANDEM REPEATS OF P-Q.
Query Match 12.4%; Score 213.5; DB 1; Length 3164;
Best Local Similarity 26.5%; Pred. No. 0.001;
Matches 91; Conservative 17; Mismatches 117; Indels 119; Gaps 15;
QY 10 RKGRLLALATAMASA-----SLTVAVPATANADPEPAPVPTTASPPSTAAAP 61
DB 2683 RHRRARSARATQASATQGMRRPALPDVTVAPTDFARP-PAPPKPPEPAPHALVSGVP 2741
QY 62 -----PAPATFVAP-PPPAANTPNAPQDPAAPPPADPNAPPPPIAPNAPQP 110
DB 2742 LPLGQAQASALPIDVPPVATGT--VLPGENRRPLTSGPAPTFRVPGVGPQR 2799
QY 111 VRIDNPGVGFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPQPPVANDRVLGRL 170
DB 2800 -RLTRP-----AVASLSRESLSPWDP---ADPTAPVLGR- 2832
QY 171 DQLYASAEATDSKAARLGSMDGMEFYMPYPTGTRINQETVSLDANGVSGSASYEVKFS 230
DB 2833 -----NPAPTSSSAG----- 2844
QY 231 PSKPNQGIWTGVISPAANAPDAGPPQRFVYVLGTA-----NNPVDKGAALAESIR 284
DB 2845 PPSPPPAV-----QVAPPTSGPPPTLTLEGVAPGPGVSRRTTRQPVAPPTTSAR 2898
QY 285 P-----LVAPPAPAPAEAPAPAPAGEVAPTPTTPO 320
DB 2899 PRGHLTVSRSLAPQPPQP-QPQPPQPQPPQP-QPQPPQPQPPQP-QPQPPQP 2940
RESULT 8
T2D3 HUMAN
ID T2D3 HUMAN STANDARD; PRT; 1083 AA.
AC 000268; Q99721; Q9BR40; Q9BX42;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Transcription initiation factor TFIID 135 kDa subunit (TAFII-135)
DE (TAFII135) (TAFII-130) (TAFII130).
GN TAF4 OR TAF4A OR TAF2C1 OR TAF2C OR TAFII135 OR TAFII130.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=97336072; PubMed=9192867;
RA Mengus G., May M., Carre L., Chambon P., Davidson I.;
RT "Human TAF(II)135 potentiates transcriptional activation by the AF-2s
RT of the retinoic acid, vitamin D3, and thyroid hormone receptors in
RT mammalian cells."
RL Genes Dev. 11:1381-1395 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.B., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehesvaisto M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.B., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20."
RL Nature 414:865-871 (2001).
RN [3]
RP SEQUENCE OF 105-1083 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=97098442; PubMed=8942982;
RA Tanese N., Saluja D., Vassallo M.F., Chen J.-L., Admon A.;
RT "Molecular cloning and analysis of two subunits of the human TFIID
RT complex: hTAFII130 and hTAFII100."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13611-13616 (1996).
RN [4]
RP IDENTIFICATION IN THE TFC-HAT COMPLEX WITH TAF5L; TAF6L; TADA3L;
RP SUPT3H; TAF2; TAF5; TERRAP; GCN5L2 AND TAF10.
RX MEDLINE=99303588; PubMed=10373431;
RA Brand M., Yamamoto K., Staub A., Iora L.;
RT "Identification of TATA-binding protein-free TAFII-containing complex
RT subunits suggests a role in nucleosome acetylation and signal
RT transduction."
RL J. Biol. Chem. 274:18285-18289 (1999).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 870-918 IN COMPLEX WITH
RP TAF12.
RX MEDLINE=20063193; PubMed=10594036;
RA Gangloff Y.-G., Werten S., Romier C., Carre L., Poch O., Moras D.,
RA Davidson I.;
RT "The human TFIID components TAF(II)135 and TAF(II)20 and the yeast
RT SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like
RT pairs."
RL Mol. Cell. Biol. 20:340-351 (2000).
CC -!- FUNCTION: Makes part of TFIID is a multimeric protein complex that
CC plays a central role in mediating promoter responses to various
CC activators and repressors. Potentiates transcriptional activation
CC by the AF-2s of the retinoic acid, vitamin D3 and thyroid hormone.
CC -!- SUBUNIT: TFIID is composed of TATA binding protein (TBP) and a
CC number of TBP-associated factors (TAFs). Component of the TFC-HAT
CC complex, at least composed of TAF5L, TAF6L, TADA3L, SUPT3H/SPT3,
CC TAF2/TAFII50, TAF4/TAFII35, TAF5/TAFII100, GCN5L2/GCN5, TAF10
CC and TERRAP.
CC -!- SUBCELLULAR LOCATION: Nuclear.
```



```
FT SEQUENCE 802 AA; 85844 MW; 592B5975EE20F77F CRC64;
SQ SEQUENCE 802 AA; 85844 MW; 592B5975EE20F77F CRC64;

Query Match 11.8%; Score 205.5; DB 1; Length 802;
Best Local Similarity 22.9%; Pred. No. 0.00072;
Matches 78; Conservative 27; Mismatches 112; Indels 123; Gaps 11;

QY 33 VPATANADPPAPVPVPTAASPPSTAAA-----PPAPATPVAPPP----- 72
DQ 343 VPRPLNKNRSPSSVNTPTSPQPPAKSCAWFTSNFSLPSPIMISSPPGKATGRPV 402
QY 73 PAAANTPNAQ-----PGDPNAA-----PPADPNAPPPVVIAPNAPQVRIQNPVG 118
DQ 403 PUCVSSVPQMPSPPTAPNGSLDSTVYVSPPTSGPAPPPPPPPPPPPPPPPPP 454
QY 119 GFSFALPAGWESDAAHFDYGSALLSKTTGDPDPGPPGPP-PPVANDTIVLGRLDQKLIAS 177
DQ 455 -----PPLPPPLPLPLAS----- 457
QY 178 AEATDSKAAARLSDMGGEFYMPYPTGTRI-----NQETVSLDANGVSGSAS---YVEVKF 228
DQ 468 -----LSHCSQASPPPTGLASTPSSKPSVLPSPSAGAPASAEITPLNPELGD 515
QY 229 SDPSKPNQIWTGIVGSPAANAPAGPPQRFVFWLGTANNPVVDKGNAAKALAESIRPLVA 288
DQ 516 SSASEP-----GLQASQPAESPTPQGLVLPAPPPPPPPPLPSGAYASALPPPG 566
QY 289 ---PPAPAPAPAPAPAPAGVAPTPTTPTPTQRTLPA 325
DQ 567 PPPPPPLPSTGPPPPPPPPPPPLNQAPPPPPPPPPAPPLPA 606

RESULT 11
TEGU_HSVBEB STANDARD; PRT; 3421 AA.
ID CPN DROME CPN DROME STANDARD; PRT; 3421 AA.
AC Q28955;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Large tegument protein.
GN 24.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1.";
RL Virology 189:304-316 (1992).
CC -!- FUNCTION: Tegument protein.
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
CC EHV-1 24, EBV BFLF1, HVS-1 64, VZV 22, AND HCMV UL48.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M86664; AB02459.1; -.
DR F1R; G36797; WZSEB6.
DR InterPro; IPR005928; Herpes_teg_N.
DR InterPro; IPR005210; Herpes_UL36.
DR Pfam; PF04843; Herpes_teg_N; 1.
DR Pfam; PF03586; Herpes_UL36; 1.
SQ SEQUENCE 3421 AA; 367078 MW; 5075EFE4739BB7AC CRC64;

Query Match 11.8%; Score 203; DB 1; Length 3421;
Best Local Similarity 23.9%; Pred. No. 0.0036;

Matches 96; Conservative 31; Mismatches 152; Indels 122; Gaps 15;

QY 6 PNLTRKGRLLAALAAAMASASLVTVAVPATANADPEPAPPVPTTAASPP-----STAAA 60
DQ 2505 PTLPPKAAAPLPSPDSASAINSGPKVFKYTFGNKSAVPPSPVAPPPTLPPAPPPLPQSTSKAAS 2564
QY 61 PPAPATPVAPPPPPAAAATPNAQGDENAAAPPADPNAP-----PPPTLPPAPLPQSTSKAASGPPPTLPPAPP 108
DQ 2565 GPPTLPPAPPLP--QSTSKAASG-----PPPTLPPAPLPQSTSKAASGPPPTLPPAPP 2617
QY 109 QVRIDNPVGGF-----SPALPAGWESDAA--- 134
DQ 2618 LPQSTSKAASGATGSDSGKTLTLDVKTQSKDKVVPVPTDKPSTTTTFAALKQSDASKPP 2677
QY 135 -----HFDYGSALLSKTTGDDP-----PFG-OPP-----PVANDTIV 166
DQ 2678 TAAIGHQKQLGTFVTPKSGDKPTDNASAPVGVSPVTPGTFGAKPFPKDAAPPVDDTK-- 2735
QY 167 LGRLDQKLIYASAEATDSKAAARLSDMGGEF-YMPYPTGTRINQETVSLDANGVSGSASYE 225
DQ 2736 ----QPVRKSLPSQVRGRPVIRPSLGPFTGTPPG-----YTIIVHGLPPSDSNVT 2783
QY 226 VKFSDPSKPNQIWTGIVGSPAANAPAGPPQRFVFWLGTANNPVVDKGNAAKALAESIRP 285
DQ 2784 QSTKEPK-----PAVETPAAPAK-----SAAAPAAAPAKSAAAPAAAP 2823
QY 286 L-VAPPPAPAPAPAPAPAPAGVAPTPTTPTPTQRTLPA 325
DQ 2824 AKSAAAPAAAPAKSAAAPAAAPAKSAAAPAAAPAKDQTKSA 2864

RESULT 12
CPN DROME CPN DROME STANDARD; PRT; 865 AA.
ID CPN DROME CPN DROME STANDARD; PRT; 865 AA.
AC Q02910;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Calphotin.
DE Calphotin.
GN CPN OR CAP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Canton-S;
RX MEDLINE=93165729; PubMed=8094559;
RA Martin J.H., Benzer S., Rudnicka M., Miller C.A.;
RT "Calphotin: a Drosophila photoreceptor cell calcium-binding protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:1531-1535(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Canton-S;
RX MEDLINE=93165730; PubMed=8434015;
RA Ballinger D.G., Xue N., Harshman K.D.;
RT "A Drosophila photoreceptor cell-specific protein, calphotin, binds
RT calcium and contains a leucine zipper.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).
CC -!- FUNCTION: Might function as a calcium-sequestering "sponge" to
CC regulate the amount of free cytoplasmic calcium. It binds 0.3 mole
CC of Ca(2+) per mole of protein.
CC -!- SUBUNIT: Homodimer (Probable).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; hypodense compartment.
CC -!- TISSUE SPECIFICITY: Soma and axons of photoreceptor cells of
CC compound eyes and ocelli.
CC -!- DEVELOPMENTAL STAGE: Expressed early in photoreceptor cell
CC development.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L02111; AAA28405.1; --
 DR EMBL; L05080; AAA28420.1; --
 DR PIR; A47282; A47282.
 DR PIR; A47283; A47283.
 DR FlyBase; Fgn0010218; Cpn.
 DR GO; GO:0005509; F:calcium ion binding; IDA.
 KW Calcium-binding.
 FT CONFLICT 36 A -> AVAPAVVA (IN REF. 2).
 FT CONFLICT 43 I -> T (IN REF. 2).
 FT CONFLICT 64 I -> V (IN REF. 2).
 FT CONFLICT 76 T -> A (IN REF. 2).
 FT CONFLICT 100 P -> PP (IN REF. 2).
 FT CONFLICT 127 VQ -> AP (IN REF. 2).
 FT CONFLICT 126 I -> V (IN REF. 2).
 FT CONFLICT 154 S -> T (IN REF. 2).
 FT CONFLICT 160 A -> E (IN REF. 2).
 FT CONFLICT 534 A -> E (IN REF. 2).
 FT CONFLICT 699 I -> T (IN REF. 2).
 FT CONFLICT 703 V -> L (IN REF. 2).
 FT CONFLICT 721 D -> E (IN REF. 2).
 SQ SEQUENCE 865 AA; 84781 NW; 2110417E0B0E7CFE CRC64;
 Query Match 11.4%; Score 197.5; DB 1; Length 865;
 Best Local Similarity 26.6%; Pred. No. 0.002;
 Matches 105; Conservative 30; Mismatches 127; Indels 133; Gaps 18;
 QY 15 LAALAIAAMASGLVTVAVPAT-----ANADP--EPAPVPTTAASPPSTAAAPP--- 62
 Db 230 VSVATKELAAEPVVVAPPATETPVVAPAAASHVSVAPAVETAVVAVPSASTEPVAA 289
 QY 63 -----APATEVAPP--AAANTPAQGDNAAPPADPNAPP----- 99
 Db 290 ATLTTAPTALAPVVAESQVAAIVVATP--PTPAPEP-ETIAPPVVAETPEVASAVA 346
 QY 100 -----PPVIAPNAPQVRIDNPGGFSFALPAGVSED-----AAHFDY 138
 Db 347 ETTTPVVPVVAESIPVAVATTV-----PATLAVTDVDTASAVPELPVPIADSPV 399
 QY 139 GSALLSKTTGD-----PPFGQPPP--VANDT-----RIVGLRQKLYASAEAT 181
 Db 400 FSA-VAETPVDLAPVLPVVAEPVPAVASEETPETPAPASAPVTIAALDIPEVAVIAA 458
 QY 182 DSKAAARLGSMDGEFYMYPGTRINQETVSLDANGVSGSAYVEVKFSDPSKNGQIWTG 241
 Db 459 PSDAPAEAPSAAPIVSTPTTASVPETAPPA-----AVTEP---IDVS 501
 QY 242 VIGSPAANADAGPPQRFVWVLTANNVDKGAALAESIRPLVAPPAPAPAPAEPA 301
 Db 502 VLSEAAIETPVA-----PPVEVITTEVAVAD-----VAPPEAAADLIIEPV 541
 QY 302 PAPA-----PAGEVAPTPTPTPTORTLP 324
 Db 542 EPPAPIPDLLEQTTSVPAVEAAESTSGSIPETSLP 576

RESULT 13

EXTN_SORBI STANDARD; PRT; 283 AA.
 AC P24152;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Extensin precursor (proline-rich glycoprotein).
 GN HRGP.
 OS Sorghum bicolor (sorghum) (Sorghum vulgare).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Sorghum.
 OX NCBI_TaxID=4558;

RN SEQUENCE FROM N.A.
 RP TISSUE=Leaf;
 RX MEDLINE=91370882; PubMed=1893107;
 RA Raz R., Cretin C., Puigdomenech P., Martinez-Izquierdo J.A.;
 RT "The sequence of a hydroxyproline-rich glycoprotein gene from Sorghum
 vulgare";
 RL Plant Mol. Biol. 16:365-367(1991).
 CC -!- FUNCTION: Structural component in primary cell wall.
 CC -!- PTM: EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE
 CC SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND THEN
 CC GLYCOSYLATED.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X56010; CAA39485.1; --
 DR PIR; S13383; S13383. Pistil extensin.
 DR InterPro; IPR003882; Pistil extensin.
 DR PRINTS; PR01218; PSTLXTENSIN.
 KW Repeat; Cell wall; Glycoprotein; Signal; Structural protein;
 FT Hydroxylation.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 283 EXTENSIN.
 SQ SEQUENCE 283 AA; 29593 MW; 8D7FCD0D8ED2D90 CRC64;
 Query Match 11.4%; Score 197; DB 1; Length 283;
 Best Local Similarity 24.0%; Pred. No. 0.00075;
 Matches 71; Conservative 17; Mismatches 108; Indels 100; Gaps 12;
 QY 34 PATANADPEPAPVPTTAASPESTAAAPAPATPVAPPPBAANTPNAQPG-DENAAAPP 92
 Db 71 PPTVTPSPKPTPP-FATPKPTPTTTPSPKPSVPYPPPPKASTPTTTPSPKPPATKPP 129
 QY 93 ADENAPPPPEVIAPNAPQPVRIIDNPGGFSFALPAGVSEDAAHFDYGSALLSKITGDPFF 152
 Db 130 TYP-TPKPPATKPP-PVYTSP-----KPPVTKEPT 160
 QY 153 PQQPPPVANDTIVIGRLDQKLYASAEATDKAAARLGSMDGEFYMYPGTRINQETVSL 212
 Db 161 PKPTPPV-----YTPNPKPPVTK----- 178
 QY 213 DANGVSGSAYVEVKFSDPSKNGQIWTGVIKSPAANAPDAGPPQRFVWVLTANNPVD 272
 Db 179 -----PPTHVSPKPTSKPTPPVT-----PSPKPKPSPP-----TYTTPKPPAT 221
 QY 273 KGAA-----KALASIRPLVAPP---PAPAPAEAPAPAPAPAGEVAPTPTTTP 319
 Db 222 KPPTSTPTPKPTPTHTPYPAHPPTYPKAPKESPPAPTPE-PTYTPPVGHTPSSPPP 276

RESULT 14

EXTN_TOBAC STANDARD; PRT; 620 AA.
 AC F13983;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Extensin precursor (Cell wall hydroxyproline-rich glycoprotein).
 GN HRGNT3.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Xanthi; TISSUE=Leaf;

15-MAR-2004 (Rel. 43, Last annotation update)
Potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 4.
HCN4.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
(1)
SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
STRAIN=Sprague-Dawley; TISSUE=Brain;
MEDLINE=94457301; PubMed=11000485;
Monteggia L.W., Eisch A.J., Tang M.D., Kaczmarek L.K., Nestler E.J.;
"Cloning and localization of the hyperpolarization-activated cyclic nucleotide-gated channel family in rat brain.";
Brain Res. Mol. Brain Res. 81:129-139(2000).
(2)
SEQUENCE OF 262-428 FROM N.A.
TISSUE=Heart;
MEDLINE=99459217; PubMed=10400919;
Shi W., Wymore R., Yu H., Wymore R.T., Pan Z., Robinson R.B., Dixon J.E., McKinnon D., Cohen I.S.;
"Distribution and prevalence of hyperpolarization-activated cation channel (HCN) mRNA expression in cardiac tissues.";
Circ. Res. 85:1-6(1999).
(3)
FUNCTION, AND TISSUE SPECIFICITY.
MEDLINE=21530492; PubMed=11675786;
Stevens D.R., Seifert R., Bufo B., Mueller F., Krenner E., Gauss R., Meyerhof W., Kaupp U.B., Lindemann B.;
"Hyperpolarization-activated channels HCN1 and HCN4 mediate responses to sour stimuli.";
Nature 413:631-635(2001).
-!- FUNCTION: Hyperpolarization-activated ion channel with very slow activation and inactivation exhibiting weak selectivity for potassium over sodium ions. May contribute to the native pacemaker currents in heart (if) and in neurons (ih). Activated by cAMP. May mediate responses to sour stimuli.
-!- SUBUNIT: The potassium channel is probably composed of a homo- or heterotetrameric complex of pore-forming subunits.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- TISSUE SPECIFICITY: Highly expressed in pyramidal and granule layer of the hippocampus, in thalamus anterior nucleus, in the supraoptic nucleus in hypothalamus, in cerebellum, and in trigeminal ganglion and superior olivary complex in the auditory system. Detected in a subset of elongated cells in taste buds.
-!- DOMAIN: The segment S4 is probably the voltage-sensor and is characterized by a series of positively charged amino acids at every third position.
-!- SIMILARITY: Belongs to the potassium channel family. HCN subfamily.
-!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
EMBL; AF247453; AAF62176.1; -
EMBL; AF155166; AAF01493.1; -
InterPro; IPR000595; CNMP_binding.
InterPro; IPR005821; Ion_trans.
InterPro; IPR005820; M-channel_nlg.
Pfam; PF00027; CNMP_binding; 1
Pfam; PF00520; ion_trans; 1
SMART; SM00100; CNMP; 1
PROSITE; PS00888; CNMP_BINDING_1; 1
PROSITE; PS00889; CNMP_BINDING_2; FALSE_NEG.
PROSITE; PS00042; CNMP_BINDING_3; 1
Transport; Ion transport; Ionic channel; Voltage-gated channel;
KW

DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Potassium/sodium hyperpolarization-activated cyclic nucleotide-gated
DE channel 4.
DE HCN4.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
OX (1)
RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RP STRAIN=Sprague-Dawley; TISSUE=Brain;
RC MEDLINE=94457301; PubMed=11000485;
RX Monteggia L.W., Eisch A.J., Tang M.D., Kaczmarek L.K., Nestler E.J.;
RA "Cloning and localization of the hyperpolarization-activated cyclic
RT nucleotide-gated channel family in rat brain.";
RL Brain Res. Mol. Brain Res. 81:129-139(2000).
RN (2)
RP SEQUENCE OF 262-428 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=99459217; PubMed=10400919;
RA Shi W., Wymore R., Yu H., Wymore R.T., Pan Z., Robinson R.B.,
Dixon J.E., McKinnon D., Cohen I.S.;
RT "Distribution and prevalence of hyperpolarization-activated cation
channel (HCN) mRNA expression in cardiac tissues.";
RL Circ. Res. 85:1-6(1999).
RN (3)
RP FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=21530492; PubMed=11675786;
RA Stevens D.R., Seifert R., Bufo B., Mueller F., Krenner E., Gauss R.,
Meyerhof W., Kaupp U.B., Lindemann B.;
RT "Hyperpolarization-activated channels HCN1 and HCN4 mediate responses
to sour stimuli.";
RL Nature 413:631-635(2001).
CC -!- FUNCTION: Hyperpolarization-activated ion channel with very slow
activation and inactivation exhibiting weak selectivity for
potassium over sodium ions. May contribute to the native pacemaker
currents in heart (if) and in neurons (ih). Activated by cAMP. May
mediate responses to sour stimuli.
CC -!- SUBUNIT: The potassium channel is probably composed of a homo- or
heterotetrameric complex of pore-forming subunits.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Highly expressed in pyramidal and granule
layer of the hippocampus, in thalamus anterior nucleus, in the
supraoptic nucleus in hypothalamus, in cerebellum, and in
trigeminal ganglion and superior olivary complex in the auditory
system. Detected in a subset of elongated cells in taste buds.
CC -!- DOMAIN: The segment S4 is probably the voltage-sensor and is
characterized by a series of positively charged amino acids at
every third position.
CC -!- SIMILARITY: Belongs to the potassium channel family. HCN
subfamily.
CC -!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; AF247453; AAF62176.1; -
EMBL; AF155166; AAF01493.1; -
InterPro; IPR000595; CNMP_binding.
InterPro; IPR005821; Ion_trans.
InterPro; IPR005820; M-channel_nlg.
Pfam; PF00027; CNMP_binding; 1
Pfam; PF00520; ion_trans; 1
SMART; SM00100; CNMP; 1
PROSITE; PS00888; CNMP_BINDING_1; 1
PROSITE; PS00889; CNMP_BINDING_2; FALSE_NEG.
PROSITE; PS00042; CNMP_BINDING_3; 1
Transport; Ion transport; Ionic channel; Voltage-gated channel;
KW

Search completed: July 7, 2004, 18:21:51
Job time : 28.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 19:18:15 ; Search time 36.5 Seconds
(without alignments)
2809.409 Million cell updates/sec

Title: US-10-720-192-2

Perfect score: 1726

Sequence: 1 MHQVDPNLTRKRLALAI.....PAGEVAPTTPTTPTQRTLPA 325

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:**

- 1: sp_archaea:**
- 2: sp_bacteria:**
- 3: sp_fungi:**
- 4: sp_human:**
- 5: sp_invertebrate:**
- 6: sp_mammal:**
- 7: sp_mhc:**
- 8: sp_organelle:**
- 9: sp_phage:**
- 10: sp_plant:**
- 11: sp_rodent:**
- 12: sp_virus:**
- 13: sp_vertebrate:**
- 14: sp_unclassified:**
- 15: sp_virus:**
- 16: sp_bacteriap:**
- 17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1539	89.2	286	2 Q3R640	Q3R640 mycobacteri
2	1053.5	61.0	368	2 Q3F4H9	Q3F4H9 mycobacteri
3	568	32.9	194	2 Q5V553	Q5V553 bovine herp
4	256	14.8	3247	12 Q6S553	Q6S553 bovine herp
5	248.5	14.4	698	10 Q9ASK4	Q9ASK4 oryza sativ
6	248	14.4	3288	12 Q7T5D9	Q7T5D9 simian herp
7	244	14.1	511	6 Q95JD0	Q95JD0 sus scrofa
8	244	14.1	565	6 Q95JD1	Q95JD1 sus scrofa
9	241.5	14.0	745	16 Q89X06	Q89X06 bradyrhizob
10	240.5	13.9	3326	12 Q7T591	Q7T591 simian herp
11	238	13.8	1269	10 Q8W5K6	Q8W5K6 oryza sativ
12	238	13.8	1269	10 Q7XH56	Q7XH56 oryza sativ
13	237.5	13.8	1188	10 Q41805	Q41805 zea mays (m
14	231	13.4	668	5 Q8MXU8	Q8MXU8 caenorhabdi
15	230.5	13.4	584	16 Q82DU7	Q82DU7 streptomyc
16	228.5	13.2	464	10 Q41645	Q41645 volvox cart

17	227.5	13.2	652	4 Q8NAA1	Q8NAA1 homo sapien
18	227	13.2	598	16 Q8VKN7	Q8VKN7 mycobacteri
19	226.5	13.1	222	10 Q9S740	Q9S740 arabidopsis
20	226.5	13.1	609	16 Q8FM53	Q8FM53 corynebacte
21	225.5	13.1	420	5 Q9VZC2	Q9VZC2 drosophila
22	225.5	13.1	652	10 Q94JZ6	Q94JZ6 arabidopsis
23	225	13.0	676	6 Q95JC9	Q95JC9 sus scrofa
24	224.5	13.0	3534	12 Q39266	Q39266 equine herp
25	223.5	12.9	652	10 Q91V48	Q91V48 arabidopsis
26	222	12.9	783	2 Q9XDH2	Q9XDH2 mycobacteri
27	222	12.9	3084	12 Q8UZ11	Q8UZ11 pseudorabie
28	218	12.6	409	10 Q9SBM1	Q9SBM1 volvox cart
29	217.5	12.6	839	16 Q9RX57	Q9RX57 deinococcus
30	217	12.6	784	13 Q90VB5	Q90VB5 gallus gall
31	217	12.6	801	5 Q23635	Q23635 caenorhabdi
32	216.5	12.5	496	16 Q7U8L8	Q7U8L8 synechococ
33	216.5	12.5	847	10 Q9XIB6	Q9XIB6 arabidopsis
34	214.5	12.4	817	3 Q07229	Q07229 saccharomyc
35	214.5	12.4	687	10 Q948Y7	Q948Y7 volvox cart
36	214.5	12.4	1016	10 Q9SPM0	Q9SPM0 zea mays (m
37	214	12.4	1204	3 Q8J0E5	Q8J0E5 varrowia li
38	214	12.4	1627	10 Q84ZL0	Q84ZL0 oryza sativ
39	213.5	12.4	760	10 Q9TOK5	Q9TOK5 arabidopsis
40	213.5	12.4	3122	12 P89459	P89459 herpes simp
41	212	12.3	753	5 Q93107	Q93107 acanthamoeb
42	211	12.2	889	16 Q9F2N5	Q9F2N5 streptomyc
43	209	12.1	1480	4 Q96Q04	Q96Q04 homo sapien
44	208.5	12.1	214	10 Q39763	Q39763 gossypium b
45	206.5	12.0	575	16 Q98F98	Q98F98 rhizobium l

ALIGNMENTS

RESULT 1

Q3R640
ID Q3R640 PRELIMINARY; PRT; 286 AA.
AC Q3R640
DT 01-MAY-2000 (TREMREL. 13, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)
DE MPT-32=45 kDa CULTURE FILTRATE glycoprotein.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP
RX MEDLINE=96196153; PubMed=8626314;
RA Dobos K.M., Khoo K.H., Swiderek K.M., Brennan P.J., Belisle J.T.;
RT "Definition of the full extent of glycosylation of the 45-kilodalton
glycoprotein of Mycobacterium tuberculosis.";
RL J. Bacteriol. 178:2498-2506(1996).
SQ SEQUENCE 286 AA; 28764 MW; 4867AD89FEB42493 CRC64;

QY	40	DPEAPPVPTTAASPPSTAAAPAPATVPAPPPAAAANTPNAQPGDPNAPPPADPNAPP	99
Db	1	DPEAPPVPTTAASPPSTAAAPAPATVPAPPPAAAANTPNAQPGDPNAPPPADPNAPP	60
QY	100	PVIAAPNAPQVRIDNPVGGFSFALPAGWESDAAHFDYGSALLSKTTGDDPPFQCPVPV	159
Db	61	PVIAAPNAPQVRIDNPVGGFSFALPAGWESDAAHFDYGSALLSKTTGDDPPFQCPVPV	120
QY	160	ANDTRIVLGRDQKLYASAEATDSKAARLGSMDGCFYMPYPTGTRINQETVSLDANGVSG	219
Db	121	ANDTRIVLGRDQKLYASAEATDSKAARLGSMDGCFYMPYPTGTRINQETVSLDANGVSG	180
QY	220	SASYTEVFPDSKNGQIWTGIVIGSPAAADAGPQRWFVWVGTTANNPVDGAAKAL	279

Query Match 89.2%; Score 1539; DB 2; Length 286;
Best Local Similarity 99.7%; Pred. No. 1.3e-89;
Matches 285; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 181 SASIYEVKFSKPKNGQIWTGVTGSPANAPDAGPPQRFVFWLGTANNPVDKGAAKAL 240
QY 280 AESIRPLVAPPAPAPAPAPAPAPAPAGEVAPTTTPTTPTQRTPLA 325
Db 241 AESIRPLVAPPAPAPAPAPAPAPAPAGEVAPTTTPTTPTQRTPLA 286

RESULT 2
Q9F4H9 PRELIMINARY; PRT; 368 AA.
AC Q9F4H9
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Fibronectin-attachment protein FAP-P (Fibronectin-binding protein p).
GN FAP.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5781;
RA Secott T.E., Lin T.L., Wu C.C.;
RT "Fibronectin-attachment protein is necessary for efficient attachment and invasion of epithelial cells by Mycobacterium avium subsp. paratuberculosis.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF395912; AAL59141.1; -
FT NON_TER 1 1
FT NON_TER 194 194
SQ SEQUENCE 194 AA; 19339 MW; 3ABE7C893DF292 CRC64;

Query Match 61.0%; Score 1053.5; DB 2; Length 368;
Best Local Similarity 60.0%; Pred. No. 6.8e-59;
Matches 228; Conservative 24; Mismatches 61; Indels 67; Gaps 12;

QY 1 MQQVDPNLTFRKRLAALIAAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
Db 1 MQQVEATSTRKGLWTLTATTGASAVAIALPATSHADPEVPTFV-----PPSTATA 54

QY 61 PPAPATPVAPPAPAA-----ANTENAO-----PGDNAAAPP--ADPNAPPPPPVIAPN 106
Db 55 PPAAAPNGQAPAPNKPAPGAPAPNGQAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 114

QY 107 APQPV-----RIDNPGVGFSPALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPPPV 159
Db 115 APPPPPADPNAGRIPNAGVGFSVLPAGWVESDASHLDYGSALLSKVTGPPMPDQPPPV 174

QY 160 ANDTRIVGLRDLQKLYASAEATDSKAAARLGSMDGFEYMPYPTGTRINQETVSLD-ANGVS 218
Db 175 ANDTRIVWGLRDLQKLYASAEANNAKAAVRLGSDMGFEYMPYPTGTRINQDSTPLANGANGST 234

QY 219 GSASYEVKFSKPKNGQIWTGVTGSPANAPDAGPPQRFVFWLGTANNPVDKGAAKA 278
Db 235 GSASYEVKFSKPKNGQIWTGVTGSPANAPDAGPPQRFVFWLGTANNPVDKGAAKA 289

QY 279 LAESIR---PLVAPPAP-----APAPAP-----APAP--A 305
Db 290 LAESIQWTFPAPAPAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAP 349

QY 306 PAGEVAPTPTTPTTPTQRTPLA 325
Db 350 PAVEVSPTP-TPTTPTQRTLSA 368

RESULT 3
Q8VS65 PRELIMINARY; PRT; 194 AA.
AC Q8VS65;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Fibronectin-attachment mutant protein (Fragment).
GN FAPP.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5781;
RA Secott T.E., Lin T.L., Wu C.C.;
RT "Fibronectin-attachment protein is necessary for efficient attachment and invasion of epithelial cells by Mycobacterium avium subsp. paratuberculosis.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF395912; AAL59141.1; -
FT NON_TER 1 1
FT NON_TER 194 194
SQ SEQUENCE 194 AA; 19339 MW; 3ABE7C893DF292 CRC64;

Query Match 32.9%; Score 568; DB 2; Length 194;
Best Local Similarity 63.9%; Pred. No. 1.3e-28;
Matches 124; Conservative 14; Mismatches 16; Indels 40; Gaps 7;

QY 162 DTRIVGLRDLQKLYASAEATDSKAAARLGSMDGFEYMPYPTGTRINQETVSLD-ANGVS 220
Db 1 DTRIVWGLRDLQKLYASAEANNAKAAVRLGSDMGFEYMPYPTGTRINQDSTPLANGANGSTCS 60

QY 221 ASYEVKFSKPKNGQIWTGVTGSPANAPDAGPPQRFVFWLGTANNPVDKGAAKALA 280
Db 61 ASYEVKFSKPKNGQIWTGVTGSPANAPDAGPPQRFVFWLGTANNPVDKGAAKALA 115

QY 281 ESIR---PLVAPPAP-----APAPAP-----APAP--A 307
Db 116 ESIAQWTFPAPAPAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAP 175

QY 308 GEVAPTPTTPTTPTQRT 321
Db 176 VEVSPTP-TPTTPTQ 188

RESULT 4
Q65553 PRELIMINARY; PRT; 3247 AA.
AC Q65553;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE UL36.
GN UL36.
OS Bovine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COOPER;
RA Schwyzer M., Vlcek C., Lowery D.E., Bello L.J., Meyer G., Misra V.,
RA Thiry E., Paces V.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-179 FROM N.A.
RC STRAIN=COOPER;
RA Schwyzer M., Styger D., Vogt B., Lowery D.E., Simard C.,
RA LaBoissiere S., Misra V., Vlcek C., Paces V.;
RL Vet. Microbiol. 0:0-0(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COOPER;

RA	Schwytzer M.;
RL	Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN	[4]
RP	SEQUENCE OF 1-179 FROM N.A.
RC	STRAIN=JURA;
RX	MEDLINE=97164286; PubMed=9010999;
RA	Schwytzer M., Styger D., Voigt B., Lowery D.E., Simard C.,
RA	Laboisiere S., Misra V., Vlcek C., Faces V.;
RT	"Gene contents in a 31-kb segment at the left genome end of bovine
RT	herpesvirus-1.";
RL	Vet. Microbiol. 53:67-77(1996).
RN	[5]
RP	SEQUENCE FROM N.A.
RC	STRAIN=JURA;
RA	Schwytzer M., Vlcek C., Lowery D.E., Bello L.J., Meyer G., Misra V.;
RL	Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN	[6]
RP	SEQUENCE FROM N.A.
RC	STRAIN=JURA;
RA	Schwytzer M.;
RL	Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; J28205; CAB01605.1; -
DR	EMBL; AJ004801; CAA06037.1; -
DR	Incerpro; IPR006928; Herpes teg.N.
DR	Incerpro; IPR005210; Herpes UL36.
DR	Pfam; PF04843; Herpes teg.N; 1.
DR	Pfam; PF03586; Herpes UL36; 1.
SQ	SEQUENCE 3247 AA; 332190 MW; 3AEAA72F8F001F6A CRC64;
Query Match 14.8%; Score 256; DB 12; Length 3247;	
Best Local Similarity 31.2%; Pred. No. 1.4e-07;	
Matches 105; Conservative 20; Mismatches 153; Indels 58; Gaps 15;	
QY	8 LTR----RKGRLAALAIAMASASIVLVAVPATANADPEAPPVPTTAAAGPSPSTAAPPA 63
Db	2584 LTRRTGPKSMFAALPWRPPAPSLVSAAVPVPGRLPAPPLPPAPLPPLPPAPLP 2643
QY	64 PATPVADP-PPAAANTNAPGDNNAAPPADNPAP-----PPVIAPNAPOQVRID 114
Db	2644 PAPELPFPAPPLPPFPAPPLPPAPPAESTAPVAPLPFPALPPALTAPTAPTAPTAPPLP 2703
QY	115 NPVGGFALPAGWVESDAAHFDYGSALLSKTTGDPPFPQP-PPVANDTRIVLGRDQK 173
Db	2704 API--TVLVAP-VPAAP-----IPAPTTPAPTAPTAPLPFPADGMANGALSATR 2754
QY	174 LYSAEATDSKAA-----RLGSDMGEPMPYPGTRINQETVSLDANGVSGSASYEVKFS 229
Db	2755 PTRRAGARKSLPAAQPSQRLLRSRASVPAPGS---ELVPPSGALGPSFV----2806
QY	230 DPSPKNGIQITVIGSPAANAADAGGPQRWFVLGTANNPVDKGAAKLAESIRLPVAP 289
Db	2807 -PSRPSSL-----EPVGLP-LPISR-----VQAPVDAPPPPAPERAPPPAP 2848
QY	290 P-PAPAPADPAPAPAPAGEVAFTPTTPTQTLP 324
Db	2849 ERAPPPAPERPPAPPPAPERAPP-----PAPERPA 2880

RESULT 5	
Q9ASK4	
Q9ASK4	
ID	Q9ASK4
AC	PRELIMINARY;
Q9ASK4	PRT; 698 AA.
OC	
Q9ASK4	
DT	01-JUN-2001 (TREMBlrel. 17, Created)
DT	01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT	01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE	Hypothetical protein.
DE	Hypothetical protein.
GN	P0439B06.8.
OS	Oryza sativa (Rice).
OS	Oryza sativa (Rice).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC	Ehrhartoideae; Oryzoideae; Oryzaeae; Oryza.
OC	NCBI_TaxID=4530;
RN	[1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone:P0439B06.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AP002862; BAE39873.1; -.
DR Gramene; Q9ASK4; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0000719; P:protein kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS01017; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 698 AA; 72733 MW; F3886E46DCA9DAE6 CRC64;
Query Match 14.4%; Score 248.5; Length 698;
Best Local Similarity 30.7%; Pred. No. 7.5e-08;
Matches 89; Conservative 20; Mismatches 134; Indels 47; Gaps 9;

QY 32 AVPATANADEPAPPVPTTAASPPSTAAAPPAPATPVAPPPPPAAATPNAPQGPNAAPP 91
DB 12 ATPTAAAPASTATPTP---ATQFNATPADPSITPPAAGPPPLFSAATPPQDPSPPSPP 68
QY 92 PADPNAPPPPVIAAPNAPOVRIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPP 151
DB 69 PAPVAVPPPATVPPPPVPVSPPT-----PSATLPPSPPPASVPVPPPTATTPPKP 121
QY 152 FPGQPPVANDTRIVLGRLDQKLYASAEATDSKAAALGSDMGFEYV-PYPGTRINQETV 210
DB 122 SPVQQPPVA-----ASPPSSPADLPPNPAPR--SDTPPVQSPSPPHRRSRPTP 170
QY 211 SLDANGVSGSAYEVKFPDPSKENGQIVGTGVGSPAANAPAGPPQRFVFWLGTANNP 270
DB 171 WAPMAPSPSGS-----PTKP-----SPASFPIAGDP-----IIPFNPP 206
QY 271 VDKGAAKALAESIRPLVAPPPAPAPAEFAPAPAPAGEVATPTTFTPTQ 320
DB 207 SSPLATPSAPSGCTPVVT-PSAPVSGPPSPGTAATAADRSNKSLSPTQ 255

RESULT 6
Q7TSD9 PRELIMINARY; PRT; 3288 AA.
AC Q7TSD9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Very large tegument protein.
GS U336.
GN Simian herpes B virus (Cercopithecoid herpesvirus 1) (Shbv).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10325;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S2489;
RX MEDLINE=22628476; PubMed=12743273;
RA Raperlygina L., Zhu L., Zurkuhlen H., Mills R., Borodovsky M.,
RA Hilliard J.K.;
RT "Complete Sequence and Comparative Analysis of the Genome of Herpes B
RT Virus (Cercopithecine Herpesvirus 1) from a Rhesus Monkey.";
RL J. Virol. 77:6167-6177(2003).
DR EMBL; AF533768; AAP1454.1; -.
SQ SEQUENCE 3288 AA; 342496 MW; 79C391EB8E6F7983 CRC64;

```

Query Match 14.4%; Score 248; DB 12; Length 3288;
Best Local Similarity 29.8%; Pred. No. 4.4e-07;
Matches 91; Conservative 24; Mismatches 150; Indels 40; Gaps 9;
QY 30 TVANPATANADPEAPPVPTTAAAGPSTAAAPPATPVAPPAAANTPNAQGDPA 89
DB 2853 TVAAAGRASAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 2911
QY 90 PPPADPNAPPV--IAFNAQPVVRIDNPGVGFSGFALPAGWVESDAAHFDYGSALLSKTT 147
DB 2912 AAFAP 2971
QY 148 GPPPPGQP-----PPVANDRIVLG-----RLDQKLYASAEATDSKAAAR-----LGS 191
DB 2972 SLTPTPSKPPAFFQPSLATGSAVAGGDFRRAPSRPTAAVPAAPSPPARLAPAVSR 3031
QY 192 DMGEFYMYPGTRINQETVSLDANGVSGSASYEVKFSDDPSKPNQIWTGVIGSPAANAP 251
DB 3032 STESALP-PDELARPTPEAPAPTETEAPVAERPAPEPPQGR-----PPSPAAP 3083
QY 252 DAGPQRFVWVWLTANNPVDKAKALAEIRPLVAPP-----PAPAPAPAPAPAPAPA 307
DB 3084 DAGP-----AAAGSGGVPAPRLGALVGRVAVPRQIPPPAPPREIPAPSPPPP 3134
QY 308 GEVAP 312
DB 3135 RSHAP 3139

RESULT 7
Q95JD0 PRELIMINARY; PRT; 511 AA.
ID Q95JD0; AC Q95JD0; DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Basic proline-rich protein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Parotid gland;
RC Zhang Q., Szalay A.A., Kyeune-Nyombi E., Sands J.F., Oberg K.C.,
RA Tieche J.-M., Leonora J.;
RT "Cloning and expression of a novel proline-rich protein from porcine
parotid glands."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035848; FAK61382.1; -
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR003882; Pistil extensin.
DR PRINTS; PR01219; PSTLEXTENSIN.
SQ SEQUENCE 511 AA; 48483 MW; AB04597964C448D7 CRC64;

Query Match 14.1%; Score 244; DB 6; Length 511;
Best Local Similarity 27.9%; Pred. No. 1e-07;
Matches 81; Conservative 10; Mismatches 99; Indels 100; Gaps 13;
QY 34 PATANADPPAPPVPTTAAAGPSTAAAPPATPVAPPAAANTPNAQGDPA 93
DB 72 PAPPGARPPGPPPP-----GPPPPGAPPGARPPGPPPP-----GPPPPGAPPGARPPPG 124
QY 94 -DPNAPPPVIAENAPQPVRIIDNPGVGFSGFALPAGWVESDAAHFDYGSALLSKTTGDDP- 151
DB 125 PPPPGPPPGAPPGARPPGPPPP-----PAGLQGGPAPSHVG-----PKKKPPPPG 172
QY 152 --FPGQPPPVANDTRIVLGRDQKLYASAEATDSKAAARLGSMDGMGFYMPYGTINQET 209
DB 173 AGHPPPPPPAN-----ESQFGPR-----PPPG----- 195
QY 210 VSLDANGVSGSASYEVKFSDDPSKPNQIWTGVIGSPAANADPAGPPQRFVWVWLTANN 269
DB 196 -----PPSPANDSQEG--SPPSADGPPQGPAP-----SGDK 225
QY 270 PVDKGAALAEIRPLVAPPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 319
DB 226 PKKKPP-----PPAGPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 265

Db 196 -----PPSPANDSQEG--SPPSADGPPQGPAP-----SGDK 225
QY 270 PVDKGAALAEIRPLVAPPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 319
DB 226 PKKKPP-----PPAGPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 265
RESULT 8
Q95JD1 PRELIMINARY; PRT; 566 AA.
ID Q95JD1; AC Q95JD1; DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Basic proline-rich protein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Parotid gland;
RC Zhang Q., Szalay A.A., Kyeune-Nyombi E., Sands J.F., Oberg K.C.,
RA Tieche J.-M., Leonora J.;
RT "Cloning and expression of a novel proline-rich protein from porcine
parotid glands."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035847; AAK61381.1; -
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR003882; Pistil extensin.
DR PRINTS; PR01219; PSTLEXTENSIN.
SQ SEQUENCE 566 AA; 53213 MW; E33B3B5E1BDEB81A CRC64;

Query Match 14.1%; Score 244; DB 6; Length 566;
Best Local Similarity 27.9%; Pred. No. 1.1e-07;
Matches 81; Conservative 10; Mismatches 99; Indels 100; Gaps 13;
QY 34 PATANADPPAPPVPTTAAAGPSTAAAPPATPVAPPAAANTPNAQGDPA 93
DB 72 PAPPGARPPGPPPP-----GPPPPGAPPGARPPGPPPP-----GPPPPGAPPGARPPPG 124
QY 94 -DPNAPPPVIAENAPQPVRIIDNPGVGFSGFALPAGWVESDAAHFDYGSALLSKTTGDDP- 151
DB 125 PPPPGPPPGAPPGARPPGPPPP-----PAGLQGGPAPSHVG-----PKKKPPPPG 172
QY 152 --FPGQPPPVANDTRIVLGRDQKLYASAEATDSKAAARLGSMDGMGFYMPYGTINQET 209
DB 173 AGHPPPPPPAN-----ESQFGPR-----PPPG----- 195
QY 210 VSLDANGVSGSASYEVKFSDDPSKPNQIWTGVIGSPAANADPAGPPQRFVWVWLTANN 269
DB 196 -----PPSPANDSQEG--SPPSADGPPQGPAP-----SGDK 225
QY 270 PVDKGAALAEIRPLVAPPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 319
DB 226 PKKKPP-----PPAGPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 265

RESULT 9
Q89X06 PRELIMINARY; PRT; 745 AA.
ID Q89X06; AC Q89X06; DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Br0521 protein.
GN BLR0521.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]

```
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RA KANEKO T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA MEDLINE=22484998; PubMed=12597275;
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005936; BAC45786.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR006665; OmpA/MotB.
DR InterPro; IPR003882; PstII_extensin.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00691; OmpA; 1.
DR PRINTS; PR01218; PSTIEXTENSIN.
DR ProDom; PD000001; Prot_kinase; 1.
KW Complete proteome.
SQ SEQUENCE 745 AA; 74545 MW; 155EDFC74DBCF6D CRC64;

Query Match 14.0%; Score 241.5; DB 16; Length 745;
Best Local Similarity 29.6%; Pred. No. 2.2e-07;
Matches 112; Conservative 16; Mismatches 136; Indels 115; Gaps 21;

Qy 15 LAALAIAM-----ASASIVTVAVPATANADPEP-----APP----- 46
Db 9 LATTALTMQLANTASHAQAGSLVVAQAQOETGPDQKFKQPPKPGGAPPAAPARPA 68
Qy 47 VPTTAAGPPSTAAPP---APATVAPPPPPAAANTPNAQGDNPAAAPPADPNAP-PPPV 102
Db 69 APPPAAPPPHPPAAPPAAAPPPAAPP---PPAARPAAPPPPPPPPAAPPKQSPPPA 126
Qy 103 IAP--NAPQPVRIINPVGFGFALPAGWVEDAHFDYGSALLSKTTGDP-PPGQPPPV 159
Db 127 AAPQOCHAPT---PPPAAPPAARPTPPAPPAAAPQAPHPPPPPPAARPTPTPPPPPA 183
Qy 160 ANDTRIVLGRLDQKLXASAEATSKAA-----ARLGSOMGEFMPYVG-----TRI 205
Db 184 GPAARPT-----PAPTATPTFPVAPPPPAAPTARPGS-----PAPATAPTTPA 227
Qy 206 NOETVSLDANGVSGSASYEVKFSKPNQIWTGVIGSPAANAPDAGPPQRFVWVLG 265
Db 228 PTATPAPTATPAGSTP-----GAPPAGRGAPPPG---VRPG 262
Qy 266 T---ANRPVDKGAALAEIRP--LVAPP-----PA--PAPAPAPAP--APAPAGEVA 311
Db 263 SPPAGSGPPAPGATPAPTTPAPGATTPPSGRGCPASTPAPGATPAPTATPAGGALT 322
Qy 312 PTP-----TPTPTORTLP 324
Db 323 PPPGRPGAGTTPGQGGTP 341

RESULT 10
ID Q7T591 PRELIMINARY; PRT; 3326 AA.
AC Q7T591;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE large tegument protein.
GN UL36.
OS Simian herpes B virus (Cercopithecine herpesvirus 1) (shbv).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_taxID=10325;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22607624; PubMed=12721804;

RA Ohsawa K., Black D.H., Sato H., Rogers K., Eberle R.;
RT "Sequence and genetic arrangement of the UL region of the monkey B
RT virus (Cercopithecine herpesvirus 1) genome and comparison with the UL
RL Arch. Virol. 148:989-997(2003).
DR EMBL; AB096160; BAC58076.2; -.
SQ SEQUENCE 3326 AA; 345566 MW; 6B53E3860F43CDF0 CRC64;

Query Match 13.9%; Score 240.5; DB 12; Length 3326;
Best Local Similarity 29.1%; Pred. No. 1.3e-06;
Matches 93; Conservative 24; Mismatches 162; Indels 41; Gaps 9;

Qy 17 ALAIAAMASASIVTVAVPATANADPEPAPVPTTAAASPESTAAAPAPATPVAP--PPPA 74
Db 2875 AAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAA 2934
Qy 75 AANTNAQGDNPAAAPPADPNAPPPV--IAPNAPQPVRIINPVGFGFALPAGWVED 132
Db 2935 AVPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAA 2994
Qy 133 AAHFDYGSALLSKTTGDPFPFGQP-----PPVANDTRIVLG-----RLDOKLYASAEATD 182
Db 2995 APAATPASPVPPTSSLPFPKPPAFQPSLATGGSVAPGGDFRRAPSRPTAAVPAAP 3054
Qy 183 SKAAAR-----LGSDMGFYMPYCTRIINQETVSLDANGVSGSASYEVKFSKPNQ 236
Db 3055 SRPPARRLARPAVSRSTESFALP-PDELARPTPEAPPTETETAPVAERPAPEPPQ 3113
Qy 237 QIWTGVIGSPAANAPDAGPPQRFVWVLGTANNPVDKGAALAEIRPLVAPP---PA 292
Db 3114 R-----PSPAAPADAGP-----AAAGSGGVPAAPRLGALVPGRAVPRQI 3157
Qy 293 PAPAPAPAPAPAPAGEVAP 312
Db 3158 PAPPREIPAPSPPPRSHAP 3177

RESULT 11
ID Q8W5K6 PRELIMINARY; PRT; 1269 AA.
AC Q8W5K6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSJNBA0079B05.10 OR OSJNAA0079B05.2.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_taxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Sasaki C., Henry D., Oates R., Simmons J.;
RT "Rice Genomic Sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079179; AAL31655.1; -.
DR EMBL; AC116601; AAM08709.1; -.
DR Gramene; Q8W5K6; -.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR008973; C2 CALB.
DR InterPro; IPR003104; FH2.
DR Pfam; PF02181; FH2; 1.
DR SMART; SM00496; FH2; 1.
KW Hypothetical protein.
SQ SEQUENCE 1269 AA; 138432 MW; 262B546481E25CA6 CRC64;
```


RESULT 14					
ID	Q8MXU8	PRELIMINARY;	PRT;	669 AA.	
AC	Q8MXU8;				
DT	01-OCT-2002	(TrEMBLrel. 22, Created)			
DT	01-OCT-2002	(TrEMBLrel. 22, last sequence update)			
DT	01-OCT-2003	(TrEMBLrel. 25, last annotation update)			
DE	Hypothetical protein K08D12.6.				
DE	K08D12.6.				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdita; Rhabditoidea;				
OC	Rhabditidae; Peleoderinae; Caenorhabditis.				
OX	NCBI_TaxID=6239;				
RN	[1]				
RX	SEQUENCE FROM N.A.				
RP	STRAIN=Bristol N2;				
RC	MEDLINE=99069613; PubMed=9851916;				
RA	Waterston R.;				
RT	"Genome sequence of the nematode C. elegans: a platform for				
RT	investigating biology. The C. elegans Sequencing Consortium.";				
RL	Science 282:2012-2018(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Bristol N2;				
RA	Waterston R.;				
RT	"The sequence of C. elegans cosmid K08D12.";				
RL	Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Bristol N2;				
RA	Waterston R.;				
RL	Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL; AC006672; AAM98004.1; --				
DR	WormPep; K08D12.6; CE21047.				
KW	Hypothetical protein.				
SQ	SEQUENCE 668 AA; 62778 MW; FB9ACC5739DC398A CRC64;				
 Query Match Best Local Similarity 13.4%; Score 231; DB 5; Length 668; Matches 107; Conservative 23; Mismatches 137; Indels 138; Gaps 17;					
QY	16	AALAIAMAGASLVTVAVATA-----NADPEAPPVPVPTAA--	----	53	
Db	169	AAAPAPADAPVEQAPVAVAPAPAAAAPDCGSAAPAAPAAATDSGYRSKNSYGD	228		
QY	54	--PPSTAAAPPA-----TPVAPPPPAAANTNAQPGD-----	85		
Db	229	QVTAPAAAAAPADAPADAPVEQAPVAVAPAPVAAPDPCEGSAAPAAAAAATDSGYRSK	288		
QY	86	-----PNAAPPADPNAPPPPVIAFNAPQPVRIIDNPVGGSFALPA-----	126		
Db	289	RNSYGDSEQVTPAAPAAAPADAPAEQAPVAVP-APAPAAAAPDCGSAAPAAAAPATD	347		
QY	127	-GWVESDAAHFDY----GSALLSKTTGDPFGOPPPVANTRIVLGRLDOKLYASA	181		
Db	348	SGYSRKNAYGDEQVTPAAAAAEPADAPV--EQAPVA-----VPAPAPTAAPAP	396		
QY	182	DKSAAARLGSDMG-----EFYMPYPCTR-----INQETVSLDANGVSGSASY	223		
Db	397	DCGSAAPAAATDSGYRSKRNAYGDEQVTPAAAAAEPADAPAEQAPVAVAPAAAA-	453		
QY	224	YEYKFSPSKNGKIWTGVISGRANAPADGPQQEWFVVWLGTANNPVDKG-AAKALA-	280		
Db	454	-----PADCGS-----AAPAAAAAPAAAP-----ATDSGYRSKRNYG	488		
QY	281	-ESIRPLVAPPAPAPAPAEAP--APAPAGEVAPTPTTTPORT	322		
Db	489	DEOVTTPAAAAAPADAPAEQAPVAVAPAPAAAAAPAAAAAPAPAAT	533		

```

Q82DUT;
01-JUN-2003 (TrEMBLrel. 24, Created)
01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
Hypotheetical protein.
SAV4871.
Streptomyces avermitilis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_TaxId=33903;
[1]
SEQUENCE FROM N.A.
STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
MEDLINE=21477403; PubMed=11572948;
Omura S., Ikeda H., Ichikawa J., Hanamoto A., Takahashi C.,
Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
Kikuchi H., Shiba I., Sakai Y., Hattori M.;
RA Kikuchi H., Shiba I., Sakai Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
avermitilis: deducing the ability of producing secondary
metabolites.";
Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
[2]
SEQUENCE FROM N.A.
STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
MEDLINE=22608306; PubMed=12692582;
Ikeda H., Ichikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
Sakai Y., Hattori M., Omura S.;
RA "Complete genome sequence and comparative analysis of the industrial
microorganism Streptomyces avermitilis.";
Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005040; BAC72583.1;
SQ Hypothetical protein; Complete proteome.
Query Match 13.4%; Score 230.5; DB 16; Length 584;
Best Local Similarity 29.6%; Pred. No. 8.4e-07;
Matches 94; Conservative 16; Mismatches 113; Indels 95; Gaps 16;

Qy 29 VTVAIPA-----TANADPPAPPVPTTAASPPSTAAAPAPATPVAPP--PPAAAN 77
Db 110 VRVGTPTAGGTGAAGGSAAPPPPPPPPP-----SPPRFVVPPVPPQEPVAGPVGPVPGD 165
Qy 78 TPNAQPGDPN-----AAPPADPNAPPPVPIAP-NAFPQVRIDNPVGGFSFALPAG 127
Db 166 APQGPPIRRDKATHLLRAVAEPTPAQWTTPQOSVAPAPAPAPVVERPASTPTHVLA- 224
Qy 128 WVESDAAHFYDGALLSKTTGDPFPFPQPPVANDRIVLGRLDOKLYASAEATDKA-- 185
Db 225 LPESAATH-----OPEQQAEPKPKVPEG-----AEAVASEAE 261
Qy 186 -AARLGSDMGFVMPYPGTRINGBTSLSDANGVSGSASYEVKFSFDPSPKNGQIWTGVIG 244
Db 262 TPVRAGTGVTFVLRAVPEQAPlTV--APGLAAL-----PGATGQ-----G 303
Qy 245 SPANAPDAGPPQRWFVWVGTTANNPNVDKGAAXALAESIRPLVAPPAPAPAPAPAPAP 304
Db 304 APAPAPGTVSPPTGAF-----GPA-----PTTATPP-PGPAPWEAPAPAP 341
Qy 305 APAGEVAPTPT---TPTP 319
Db 342 EPAPAPAPAQGTPTTP 359

Search completed: July 7, 2004, 18:23:02
Job time : 39.5 secs

```

Search completed: July 7, 2004, 18:23:02
Job time : 39.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 18:21:10 ; Search time 16.5 Seconds
(without alignments)
1016.875 Million cell updates/sec

Title: US-10-720-192-2
Perfect score: 1726
Sequence: 1 MHQVDNLTTRKGRLAALAI.....PAGEVAPTPTPTPTQTLPA 325

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1726	100.0	325	1	US-08-382-184-2
2	1726	100.0	325	2	US-08-641-356-2
3	1726	100.0	325	3	US-09-132-528-2
4	1726	100.0	325	3	US-09-132-528-3
5	1726	100.0	325	3	US-08-875-494-2
6	1726	100.0	325	4	US-08-599-366-2
7	1726	100.0	325	4	US-08-599-366-3
8	1726	100.0	325	4	US-08-875-494-2
9	1726	100.0	325	4	US-09-510-031A-6
10	1726	100.0	325	4	US-09-985-372-2
11	1726	100.0	325	4	US-09-985-372-3
12	1726	100.0	332	4	US-08-818-112-53
13	1726	100.0	332	4	US-08-818-111-53
14	1726	100.0	332	4	US-09-086-556-53
15	1726	100.0	332	4	US-09-072-596-53
16	1726	100.0	332	4	US-09-072-967-53
17	1720	99.7	325	4	US-09-510-031A-2
18	1542	89.3	286	1	US-08-382-184-3
19	1542	89.3	286	2	US-08-641-356-3
20	1542	89.3	286	3	US-08-132-528-4
21	1542	89.3	286	3	US-08-875-494-3
22	1542	89.3	286	4	US-09-599-366-4
23	1542	89.3	286	4	US-08-875-494-3
24	1542	89.3	286	4	US-09-985-372-4
25	1539	89.2	285	4	US-09-510-031A-7
26	1499	86.8	652	4	US-09-072-596-350
27	1499	86.8	652	4	US-09-072-967-355

28	1499	86.8	802	4	US-09-056-556-214	Sequence 214, App
29	1499	86.8	802	4	US-09-072-596-209	Sequence 209, App
30	1499	86.8	802	4	US-09-072-596-346	Sequence 346, App
31	1499	86.8	802	4	US-09-072-967-214	Sequence 214, App
32	1499	86.8	802	4	US-09-072-967-351	Sequence 351, App
33	1499	86.8	802	4	US-09-287-849-10	Sequence 10, Appl
34	985	57.1	381	4	US-09-510-031A-5	Sequence 5, Appl
35	980	56.8	287	4	US-08-311-731A-96	Sequence 96, Appl
36	980	56.8	287	4	US-09-510-031A-4	Sequence 4, Appl
37	969	56.1	287	4	US-09-510-031A-8	Sequence 8, Appl
38	233	13.5	8991	4	US-08-714-741-32	Sequence 32, Appl
39	200.5	11.6	214	1	US-08-217-327-4	Sequence 4, Appl
40	200	11.6	408	1	US-07-609-716-65	Sequence 65, Appl
41	200	11.6	408	3	US-08-475-411A-65	Sequence 65, Appl
42	200	11.6	408	3	US-08-478-029A-65	Sequence 65, Appl
43	197.5	11.4	330	1	US-08-642-255-32	Sequence 32, Appl
44	197	11.4	538	4	US-09-616-288-43	Sequence 43, Appl
45	196.5	11.4	666	4	US-09-050-739-70	Sequence 70, Appl

ALIGNMENTS

RESULT 1
US-08-382-184-2
; Sequence 2, Application US/08382184
; Patent No. 5714593
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: MICROBACTERIAL PROTEINS.
; TITLE OF INVENTION: MICROORGANISMS PRODUCING THEM AND THEIR USE FOR VACCINES
; TITLE OF INVENTION: AND FOR THE DETECTION OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER &
; ADDRESSEE: NEUSTADT, P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION NUMBER: US/08/382,184
; FILING DATE: 01-FEB-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 22640720
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-382-184-2

Query Match 100.0%; Score 1726; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MHQVDNLTTRKGRLAALAIAMASASLVTVAVPATANADPEPVPVPTTASPPSTAAA 60
Db 1 MHQVDNLTTRKGRLAALAIAMASASLVTVAVPATANADPEPVPVPTTASPPSTAAA 60


```
Db 301 APAPAGEVAPTPTTPTQRTLPA 325

RESULT 4
US-09-132-528-3
; Sequence 3, Application US/09132528A
; Patent No. 6221353
; GENERAL INFORMATION:
; APPLICANT: LAQUEYERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 0660-0138-0D1V
; CURRENT APPLICATION NUMBER: US/09/132.528A
; CURRENT FILING DATE: 1998-08-11
; EARLIER APPLICATION NUMBER: 08/641.356
; EARLIER FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-132-528-3

Query Match 100.0%; Score 1726; DB 3; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDPNLTRKGRLLAALAAASASLVTVAVPATANADPEAPVPVTTAASPSTAAA 60
Db 1 MHQVDPNLTRKGRLLAALAAASASLVTVAVPATANADPEAPVPVTTAASPSTAAA 60
Qy 61 PPAPATVPAPPPAAAANTPNACQDPAADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
Db 61 PPAPATVPAPPPAAAANTPNACQDPAADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRIVLGRDQKLYASAEA 180
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRIVLGRDQKLYASAEA 180
Qy 181 TDSKAAARLGSDMGEFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDFSPKNGQIWT 240
Db 181 TDSKAAARLGSDMGEFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDFSPKNGQIWT 240
Qy 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAESIRPLVAPPAPAPAPAE 300
Db 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAESIRPLVAPPAPAPAE 300
Qy 301 APAPAGEVAPTPTTPTQRTLPA 325
Db 301 APAPAGEVAPTPTTPTQRTLPA 325

RESULT 6
US-09-599-366-2
; Sequence 2, Application US/09599366
; Patent No. 6335181
; GENERAL INFORMATION:
; APPLICANT: LAQUEYERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 0660-0138-0D1V
; CURRENT APPLICATION NUMBER: US/09/599.366
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 09/132.528
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 08/641.356
; PRIOR FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-599-366-2

Query Match 100.0%; Score 1726; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDPNLTRKGRLLAALAAASASLVTVAVPATANADPEAPVPVTTAASPSTAAA 60
Db 1 MHQVDPNLTRKGRLLAALAAASASLVTVAVPATANADPEAPVPVTTAASPSTAAA 60
Qy 61 PPAPATVPAPPPAAAANTPNACQDPAADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
Db 61 PPAPATVPAPPPAAAANTPNACQDPAADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRIVLGRDQKLYASAEA 180
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGQPPVANDTRIVLGRDQKLYASAEA 180
Qy 181 TDSKAAARLGSDMGEFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDFSPKNGQIWT 240
Db 181 TDSKAAARLGSDMGEFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDFSPKNGQIWT 240
Qy 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAESIRPLVAPPAPAPAE 300
Db 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAESIRPLVAPPAPAE 300
Qy 301 APAPAGEVAPTPTTPTQRTLPA 325
Db 301 APAPAGEVAPTPTTPTQRTLPA 325

RESULT 5
US-08-875-494-2
; Sequence 2, Application US/08875494
; Patent No. 6221610
; GENERAL INFORMATION:
; APPLICANT: LAQUEYERIE, ANNE
; APPLICANT: MARCHAL, GILLES
; APPLICANT: PESCHER, PASCALE
; APPLICANT: ROMAIN, FELIX
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USES FOR VACCINES AND FOR THE DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 0860-0122-0 PCT
; CURRENT APPLICATION NUMBER: US/08/875.494
; CURRENT FILING DATE: 1997-08-01
; EARLIER APPLICATION NUMBER: PCT/FR96/00166
```

Db 1 MHQVDNLTTRKGRLLAALAAIAMAASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
QY 61 PPAPATPVAPPAAANTPNAQPGDPAAPADPNAPPPPPVIAPNAPQVRIDNPVGGF 120
Db 61 PPAPATPVAPPAAANTPNAQPGDPAAPADPNAPPPPPVIAPNAPQVRIDNPVGGF 120
QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPPQPPVANDTRIVLGRDQKLYASAEA 180
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPPQPPVANDTRIVLGRDQKLYASAEA 180
QY 181 TDSKAAARLGSDMGDFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDFSKNGQIWT 240
Db 181 TDSKAAARLGSDMGDFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDFSKNGQIWT 240
QY 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAESIRPLVAPPAPAPAPAE 300
Db 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAESIRPLVAPPAPAPAPAE 300
QY 301 APAPAPAGEVAPTPTTPTPQRTLPA 325
Db 301 APAPAPAGEVAPTPTTPTPQRTLPA 325

RESULT 7
US-09-599-366-3
; Sequence 3, Application US/09599366
; Patent No. 6335181
; GENERAL INFORMATION:
; APPLICANT: LAQUEYERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/599,366
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 09/132,528
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 08/641,356
; PRIOR FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-599-366-3

Query Match 100.0%; Score 1726; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127; Mismatches 0; Indels 0; Gaps 0;
Matches 325; Conservative 0;
QY 1 MHQVDNLTTRKGRLLAALAAIAMAASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
Db 1 MHQVDNLTTRKGRLLAALAAIAMAASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
QY 61 PPAPATPVAPPAAANTPNAQPGDPAAPADPNAPPPPPVIAPNAPQVRIDNPVGGF 120
Db 61 PPAPATPVAPPAAANTPNAQPGDPAAPADPNAPPPPPVIAPNAPQVRIDNPVGGF 120
QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPPQPPVANDTRIVLGRDQKLYASAEA 180
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPPQPPVANDTRIVLGRDQKLYASAEA 180
QY 181 TDSKAAARLGSDMGDFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDFSKNGQIWT 240
Db 181 TDSKAAARLGSDMGDFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDFSKNGQIWT 240
QY 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAESIRPLVAPPAPAPAPAE 300
Db 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAESIRPLVAPPAPAPAPAE 300

QY 301 APAPAPAGEVAPTPTTPTPQRTLPA 325
Db 301 APAPAPAGEVAPTPTTPTPQRTLPA 325
RESULT 8
US-08-875-494-2
; Sequence 2, Application US/08875494
; Patent No. 6379902
; GENERAL INFORMATION:
; APPLICANT: LAQUEYERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USES FOR VACCINES AND FOR THE DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 0660-0122-0 PCT
; CURRENT APPLICATION NUMBER: US/08/875,494
; CURRENT FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: PCT/FR96/00166
; PRIOR FILING DATE: 1996-02-01
; PRIOR APPLICATION NUMBER: 382184
; PRIOR FILING DATE: 1995-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-08-875-494-2

Query Match 100.0%; Score 1726; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127; Mismatches 0; Indels 0; Gaps 0;
Matches 325; Conservative 0;
QY 1 MHQVDNLTTRKGRLLAALAAIAMAASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
Db 1 MHQVDNLTTRKGRLLAALAAIAMAASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
QY 61 PPAPATPVAPPAAANTPNAQPGDPAAPADPNAPPPPPVIAPNAPQVRIDNPVGGF 120
Db 61 PPAPATPVAPPAAANTPNAQPGDPAAPADPNAPPPPPVIAPNAPQVRIDNPVGGF 120
QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPPQPPVANDTRIVLGRDQKLYASAEA 180
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPPQPPVANDTRIVLGRDQKLYASAEA 180
QY 181 TDSKAAARLGSDMGDFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDFSKNGQIWT 240
Db 181 TDSKAAARLGSDMGDFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDFSKNGQIWT 240
QY 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAESIRPLVAPPAPAPAPAE 300
Db 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAESIRPLVAPPAPAPAPAE 300
QY 301 APAPAPAGEVAPTPTTPTPQRTLPA 325
Db 301 APAPAPAGEVAPTPTTPTPQRTLPA 325

RESULT 9
US-09-510-031A-6
; Sequence 6, Application US/09510031A
; Patent No. 6638518
; GENERAL INFORMATION:
; APPLICANT: Ratliff, Timothy
; APPLICANT: Kline, Joel
; TITLE OF INVENTION: METHOD FOR INHIBITING INFLAMMATORY RESPONSES
; FILE REFERENCE: 140.0010 0101
; CURRENT APPLICATION NUMBER: US/09/510,031A
; CURRENT FILING DATE: 2000-02-22

```
; PRIOR APPLICATION NUMBER: US 60/121,177
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-510-031A-6

Query Match          100.0%; Score 1726; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDPNLTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPPVPTTAAASPSTAAA 60
DB 1 MHQVDPNLTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPPVPTTAAASPSTAAA 60

QY 61 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
DB 61 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120

QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPVANDTRIVLGRDQKLYASAEA 180
DB 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPVANDTRIVLGRDQKLYASAEA 180

QY 181 TDSKAAARLGSDMGCEFYMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 240
DB 181 TDSKAAARLGSDMGCEFYMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 240

QY 241 GVIGSPAANADAGPPQRFVWLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300
DB 241 GVIGSPAANADAGPPQRFVWLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300

QY 301 APAPAPAGEVAPTTPTTPTQRTLPA 325
DB 301 APAPAPAGEVAPTTPTTPTQRTLPA 325

RESULT 11
US-09-985-372-3
; Sequence 3, Application US/09985372
; Patent No. 6676945
; GENERAL INFORMATION:
; APPLICANT: LAQUEYERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/985,372
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/132,528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-985-372-3

Query Match          100.0%; Score 1726; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDPNLTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPPVPTTAAASPSTAAA 60
DB 1 MHQVDPNLTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPPVPTTAAASPSTAAA 60

QY 61 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
DB 61 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120

QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPVANDTRIVLGRDQKLYASAEA 180
DB 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPVANDTRIVLGRDQKLYASAEA 180

QY 181 TDSKAAARLGSDMGCEFYMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 240
DB 181 TDSKAAARLGSDMGCEFYMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 240

QY 241 GVIGSPAANADAGPPQRFVWLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300
DB 241 GVIGSPAANADAGPPQRFVWLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300

QY 301 APAPAPAGEVAPTTPTTPTQRTLPA 325
DB 301 APAPAPAGEVAPTTPTTPTQRTLPA 325

RESULT 10
US-09-985-372-2
; Sequence 2, Application US/09985372
; Patent No. 6676945
; GENERAL INFORMATION:
; APPLICANT: LAQUEYERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/985,372
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/132,528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-985-372-2

Query Match          100.0%; Score 1726; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDPNLTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPPVPTTAAASPSTAAA 60
DB 1 MHQVDPNLTRKGRGLAALAIAMASASLVTVAVPATANADPEPAPPVPTTAAASPSTAAA 60

QY 61 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
```



```

Db 308 APAPAPAGEVAPTPTTPTQRTLPA 332
|||||
RESULT 14
US-09-056-556-53
; Sequence 53, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
; NUMBER OF SEQUENCES: 241
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,556
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 622-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-056-556-53
Query Match 100.0%; Score 1726; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 6.5e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDPNLTTRKGRLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
Db 8 MHQVDPNLTTRKGRLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 67
Qy 61 PPAPATPVAPPPAAANTPNAQGDPAAPADPNAPPPVPIAPNAPQVRIDNPVGGF 120
Db 68 PPAPATPVAPPPAAANTPNAQGDPAAPADPNAPPPVPIAPNAPQVRIDNPVGGF 127
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPQPPVANDTRIVLGRLDQKLYASAEA 180
Db 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPQPPVANDTRIVLGRLDQKLYASAEA 187
Qy 181 TDSKAAARLGSMDGEFYMPYPGTRINQETVSLDANGVSGSASYEYVKFSPKPNQGIWT 240
Db 188 TDSKAAARLGSMDGEFYMPYPGTRINQETVSLDANGVSGSASYEYVKFSPKPNQGIWT 247
Qy 241 GVIGSPAANADAGPPQRFVWVLGTANNPVVDKGAALAEISIRPLVAPPAPAPAPAP 300
Db 248 GVIGSPAANADAGPPQRFVWVLGTANNPVVDKGAALAEISIRPLVAPPAPAPAPAP 307
Qy 301 APAPAPAGEVAPTPTTPTQRTLPA 325
Db 308 APAPAPAGEVAPTPTTPTQRTLPA 332

```

```

RESULT 15
US-09-072-596-53
; Sequence 53, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-072-596-53
Query Match 100.0%; Score 1726; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 6.5e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDPNLTTRKGRLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
Db 8 MHQVDPNLTTRKGRLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 67
Qy 61 PPAPATPVAPPPAAANTPNAQGDPAAPADPNAPPPVPIAPNAPQVRIDNPVGGF 120
Db 68 PPAPATPVAPPPAAANTPNAQGDPAAPADPNAPPPVPIAPNAPQVRIDNPVGGF 127
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPQPPVANDTRIVLGRLDQKLYASAEA 180
Db 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPQPPVANDTRIVLGRLDQKLYASAEA 187
Qy 181 TDSKAAARLGSMDGEFYMPYPGTRINQETVSLDANGVSGSASYEYVKFSPKPNQGIWT 240
Db 188 TDSKAAARLGSMDGEFYMPYPGTRINQETVSLDANGVSGSASYEYVKFSPKPNQGIWT 247
Qy 241 GVIGSPAANADAGPPQRFVWVLGTANNPVVDKGAALAEISIRPLVAPPAPAPAPAP 300
Db 248 GVIGSPAANADAGPPQRFVWVLGTANNPVVDKGAALAEISIRPLVAPPAPAPAPAP 307
Qy 301 APAPAPAGEVAPTPTTPTQRTLPA 325

```

Thu Jul 8 11:11:29 2004

us-10-720-192-2.ra1

Page 8

Db 308 APAPAPAGEVAPTPTTPTQRTLPA 332
|||||

Search completed: July 7, 2004, 18:24:52
Job time : 25.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 18:24:06 ; Search time 42.5 Seconds
(without alignments)
2380.406 Million cell updates/sec

Title: US-10-720-192-2

Perfect score: 1726

Sequence: 1 MHQVDPNLTTRKGRLLAALAI.....PAGEVAPPTTPPTQRTLPA 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1726	100.0	325	10	US-09-985-372-2 Sequence 2, Appl 1
2	1726	100.0	325	10	US-09-985-372-3 Sequence 3, Appl 1
3	1726	100.0	332	12	US-09-886-349A-41 Sequence 41, Appl 1
4	1726	100.0	332	14	US-10-193-002-53 Sequence 53, Appl 1
5	1726	100.0	332	14	US-10-084-843-53 Sequence 53, Appl 1
6	1726	100.0	332	14	US-10-098-732A-41 Sequence 41, Appl 1
7	1542	89.3	286	10	US-09-985-372-4 Sequence 4, Appl 1
8	1499	86.8	652	14	US-10-193-002-350 Sequence 350, App
9	1499	86.8	652	14	US-10-084-843-355 Sequence 355, App
10	1499	86.8	802	9	US-09-287-849-10 Sequence 10, Appl 1
11	1499	86.8	802	14	US-10-193-002-209 Sequence 209, App
12	1499	86.8	802	14	US-10-193-002-346 Sequence 346, App
13	1499	86.8	802	14	US-10-084-843-214 Sequence 214, App
14	1499	86.8	802	14	US-10-084-843-351 Sequence 351, App
15	1499	86.8	802	14	US-10-359-460-10 Sequence 10, Appl 1

16	248.5	14.4	273	16	US-10-437-963-184760	Sequence 184760,
17	241	14.0	402	16	US-10-437-963-141313	Sequence 141313,
18	240.5	13.9	437	16	US-10-437-963-168762	Sequence 168762,
19	239	13.8	431	16	US-10-437-963-204963	Sequence 204963,
20	238	13.8	1269	16	US-10-437-963-190985	Sequence 190985,
21	237.5	13.8	597	12	US-10-425-114-41545	Sequence 41545, A
22	237.5	13.5	547	16	US-10-437-963-102579	Sequence 102579,
23	232.5	13.5	270	16	US-10-437-963-130068	Sequence 130068,
24	232.5	13.5	360	16	US-10-437-963-179473	Sequence 179473,
25	230.5	13.4	426	16	US-10-437-963-142229	Sequence 142229,
26	230.5	13.4	584	14	US-10-156-761-12405	Sequence 12405, A
27	229.5	13.3	569	16	US-10-437-963-149276	Sequence 149276,
28	229	13.3	440	16	US-10-437-963-128104	Sequence 128104,
29	228.5	13.2	235	16	US-10-437-963-162861	Sequence 162861,
30	227.5	13.2	652	15	US-10-104-047-3364	Sequence 3364, Ap
31	227	13.2	598	12	US-10-282-122A-64954	Sequence 64954, A
32	226.5	13.1	504	16	US-10-437-963-128105	Sequence 128105,
33	223.5	12.9	299	16	US-10-437-963-134054	Sequence 134054,
34	223.5	12.9	501	16	US-10-437-963-161137	Sequence 161137,
35	223.5	12.9	533	16	US-10-437-963-167716	Sequence 167716,
36	221.5	12.8	556	16	US-10-437-963-137918	Sequence 137918,
37	220.5	12.8	406	16	US-10-437-963-122256	Sequence 122256,
38	220	12.7	389	12	US-10-425-114-56400	Sequence 56400, A
39	220	12.7	466	16	US-10-437-963-195119	Sequence 195119,
40	219.5	12.7	369	16	US-10-437-963-196554	Sequence 196554,
41	219.5	12.7	376	16	US-10-437-963-154970	Sequence 154970,
42	218.5	12.7	285	16	US-10-437-963-189482	Sequence 189482,
43	218.5	12.7	534	16	US-10-437-963-159895	Sequence 159895,
44	217.5	12.6	350	16	US-10-437-963-102583	Sequence 102583,
45	217.5	12.6	817	15	US-10-369-493-1813	Sequence 1813, Ap

ALIGNMENTS

RESULT 1

US-09-985-372-2
; Sequence 2, Application US/09985372
; Publication No. US20030054008A1
; GENERAL INFORMATION:
; APPLICANT: LAQUEYERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/985,372
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/132,528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-985-372-2

Query Match 100.0%; Score 1726; DB 10; Length 325;
Best Local Similarity 100.0%; Pred. No. 8.6e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MHQVDPNLTTRKGRLLAALAI	AAMASASLVTVAVPATANADPEPPVPTTAA	SPSTAAA	60
Db	1	MHQVDPNLTTRKGRLLAALAI	AAMASASLVTVAVPATANADPEPPVPTTAA	SPSTAAA	60
QY	61	PPAPATPVAPPPAAANTNAQ	GGPNAAPPADNAPPPVIA	NPQVRI	120
Db	61	PPAPATPVAPPPAAANTNAQ	GGPNAAPPADNAPPPVIA	NPQVRI	120
QY	121	SPALPAGWVESDAAHFYDGS	ALLSKTTGTPPPPGPPVAND	TRIVLGR	180

```

; Publication No. US20040086523A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Reed, Steven
; APPLICANT: Alderson, Mark
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009070US
; CURRENT APPLICATION NUMBER: US/09/886,349A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/597,796
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 60/265,737
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: DPEP
US-09-886-349A-41

```

```

Query Match 100.0%; Score 1726; DB 12; Length 332;
Best Local Similarity 100.0%; Pred. No. 8.7e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEAPVPTTAAASPFSTAAA 60
DB 8 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEAPVPTTAAASPFSTAAA 67
QY 61 PPAPATPVAPPPAAANTPNAQPGDPNAPADPNAPPPVIAENAPQVRIDNPVGGF 120
DB 68 PPAPATPVAPPPAAANTPNAQPGDPNAPADPNAPPPVIAENAPQVRIDNPVGGF 127
QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGPPVANDTRIVLGRDQKLYASAEA 180
DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGPPVANDTRIVLGRDQKLYASAEA 187
QY 181 TDSKAAARLGSDMGDFYMPYFGTRINQETVSLDANGVSGSASYEYKFSKPNQGIWT 240
DB 188 TDSKAAARLGSDMGDFYMPYFGTRINQETVSLDANGVSGSASYEYKFSKPNQGIWT 247
QY 241 GVIGSPAANAPDAGPPQRFVWLGTANNPVVDKGAALAESIRPLVAPPPAPAPAPAE 300
DB 248 GVIGSPAANAPDAGPPQRFVWLGTANNPVVDKGAALAESIRPLVAPPPAPAPAPAE 307
QY 301 APAPAGEVAPTPTTPTPQRTLPA 325
DB 308 APAPAGEVAPTPTTPTPQRTLPA 332

```

```

RESULT 4
US-10-193-002-53
; Sequence 53, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, David C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vegdick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue

```

```

; Publication No. US20030054008A1
; GENERAL INFORMATION:
; APPLICANT: LAQUEYRERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/985,372
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/132,528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-985-372-3

```

```

Query Match 100.0%; Score 1726; DB 10; Length 325;
Best Local Similarity 100.0%; Pred. No. 8.6e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEAPVPTTAAASPFSTAAA 60
DB 1 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEAPVPTTAAASPFSTAAA 60
QY 61 PPAPATPVAPPPAAANTPNAQPGDPNAPADPNAPPPVIAENAPQVRIDNPVGGF 120
DB 61 PPAPATPVAPPPAAANTPNAQPGDPNAPADPNAPPPVIAENAPQVRIDNPVGGF 120
QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGPPVANDTRIVLGRDQKLYASAEA 180
DB 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGPPVANDTRIVLGRDQKLYASAEA 180
QY 181 TDSKAAARLGSDMGDFYMPYFGTRINQETVSLDANGVSGSASYEYKFSKPNQGIWT 240
DB 181 TDSKAAARLGSDMGDFYMPYFGTRINQETVSLDANGVSGSASYEYKFSKPNQGIWT 240
QY 241 GVIGSPAANAPDAGPPQRFVWLGTANNPVVDKGAALAESIRPLVAPPPAPAPAPAE 300
DB 241 GVIGSPAANAPDAGPPQRFVWLGTANNPVVDKGAALAESIRPLVAPPPAPAPAPAE 300
QY 301 APAPAGEVAPTPTTPTPQRTLPA 325
DB 301 APAPAGEVAPTPTTPTPQRTLPA 325

```

```

RESULT 3
US-09-886-349A-41
; Sequence 41, Application US/09886349A

```

;
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/193,002
; FILING DATE: 10-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-10-193-002-53

Query Match 100.0%; Score 1726; DB 14; Length 332;
Best Local Similarity 100.0%; Pred. No. 8.7e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHQVDNLTREKRLAALAIAMASASLVTVAVPATANADPPAPPVPTTAASPPSTAAA 60
DB 8 MHQVDNLTREKRLAALAIAMASASLVTVAVPATANADPPAPPVPTTAASPPSTAAA 67
QY 61 PPAPATPVAPPVPPAAANTPNAQPGDPAADPPADNAPPPVIAAPNAPQPVRIIDNPVGGF 120
DB 68 PPAPATPVAPPVPPAAANTPNAQPGDPAADPPADNAPPPVIAAPNAPQPVRIIDNPVGGF 127
QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPPPVANDTRIVLGRDQKLYASAEA 180
DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPPPVANDTRIVLGRDQKLYASAEA 187
QY 181 TDSKAAARLGSMDGEFYMPYEGTRINQETVSLDANGVSGSAYYEVKFSPPSKNGQIWT 240
DB 188 TDSKAAARLGSMDGEFYMPYEGTRINQETVSLDANGVSGSAYYEVKFSPPSKNGQIWT 247
QY 241 GVIGSPAANAPDAGPPQRFVFWLGTANNPVDKGAALKALAESIRPLVAPPPAPAPAPAE 300
DB 248 GVIGSPAANAPDAGPPQRFVFWLGTANNPVDKGAALKALAESIRPLVAPPPAPAPAPAE 307
QY 301 APAPAPAGEVAPTPTTPTPQRTLPA 325
DB 308 APAPAPAGEVAPTPTTPTPQRTLPA 332

RESULT 5
US-10-084-843-53
; Sequence 53, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Skeiky, Yasir A.W.
; Dillon, David C.
; Campos-Neto, Antonio
; Houghton, Raymond
; Vedwick, Thomas S.
; Twardzik, Daniel R.

;
; Lodes, Michael J.
; Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SRED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-10-084-843-53

Query Match 100.0%; Score 1726; DB 14; Length 332;
Best Local Similarity 100.0%; Pred. No. 8.7e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHQVDNLTREKRLAALAIAMASASLVTVAVPATANADPPAPPVPTTAASPPSTAAA 60
DB 8 MHQVDNLTREKRLAALAIAMASASLVTVAVPATANADPPAPPVPTTAASPPSTAAA 67
QY 61 PPAPATPVAPPVPPAAANTPNAQPGDPAADPPADNAPPPVIAAPNAPQPVRIIDNPVGGF 120
DB 68 PPAPATPVAPPVPPAAANTPNAQPGDPAADPPADNAPPPVIAAPNAPQPVRIIDNPVGGF 127
QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPPPVANDTRIVLGRDQKLYASAEA 180
DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPPPVANDTRIVLGRDQKLYASAEA 187
QY 181 TDSKAAARLGSMDGEFYMPYEGTRINQETVSLDANGVSGSAYYEVKFSPPSKNGQIWT 240
DB 188 TDSKAAARLGSMDGEFYMPYEGTRINQETVSLDANGVSGSAYYEVKFSPPSKNGQIWT 247
QY 241 GVIGSPAANAPDAGPPQRFVFWLGTANNPVDKGAALKALAESIRPLVAPPPAPAPAPAE 300
DB 248 GVIGSPAANAPDAGPPQRFVFWLGTANNPVDKGAALKALAESIRPLVAPPPAPAPAPAE 307
QY 301 APAPAPAGEVAPTPTTPTPQRTLPA 325
DB 308 APAPAPAGEVAPTPTTPTPQRTLPA 332

RESULT 6
US-10-098-732A-41
; Sequence 41, Application US/10098732A
; Publication No. US20030175294A1

```

; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Gudarian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:DPEP
US-10-098-732A-41

Query Match      100.0%; Score 1726; DB 14; Length 332;
Best Local Similarity 100.0%; Pred. No. 8.7e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQVDPNLTERRKRLAALAIAMASASLVTVAVPATANADPEAPPVPTTAASPPSTAAA 60
DB 8 MEQVDPNLTERRKRLAALAIAMASASLVTVAVPATANADPEAPPVPTTAASPPSTAAA 67

QY 61 PPAPATPVAPPPAAANTPNAQGDENAPADPNAPPPVVIAPNAPQPVRIINPVGGF 120
DB 68 PPAPATPVAPPPAAANTPNAQGDENAPADPNAPPPVVIAPNAPQPVRIINPVGGF 127

QY 121 SPALPAGWVESDAARHFDYGSALLSKTTGDPFGGPPGVANDTRIVLGRDQKLYASAEA 180
DB 128 SPALPAGWVESDAARHFDYGSALLSKTTGDPFGGPPGVANDTRIVLGRDQKLYASAEA 187

QY 181 TDSKAAARLGSDMGDFYFYPGTRINQETVSLDANGVSGSASYEVKFSDFPSKNGQIWT 240
DB 188 TDSKAAARLGSDMGDFYFYPGTRINQETVSLDANGVSGSASYEVKFSDFPSKNGQIWT 247

QY 241 GVIGSPAANAPDAGPPQRFVFWLGTANNPVDKGAALASIRPLVAPPAPAPAPAP 300
DB 248 GVIGSPAANAPDAGPPQRFVFWLGTANNPVDKGAALASIRPLVAPPAPAPAPAP 307

QY 301 APAPAPAGEVAPTPTTPTPORTLPA 325
DB 308 APAPAPAGEVAPTPTTPTPORTLPA 332

RESULT 7
US-09-985-372-4
; Sequence 4, Application US/09985372
; Publication No. US20030054008A1
; GENERAL INFORMATION:
; APPLICANT: LAQUERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICRORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/985,372
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/132,528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 286
; TYPE: PRT

```

```

; ORGANISM: Mycobacterium tuberculosis
US-09-985-372-4

Query Match      89.3%; Score 1542; DB 10; Length 286;
Best Local Similarity 100.0%; Pred. No. 9.3e-94;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 DPEPAPPVPTTAASPPSTAAAPAPATPVAPPVPPAAANTPNAQGDENAPADPNAPP 99
DB 1 DPEPAPPVPTTAASPPSTAAAPAPATPVAPPVPPAAANTPNAQGDENAPADPNAPP 60

QY 100 PVIAENAPQPVRIINPVGGFSAFALPAGWVESDAARHFDYGSALLSKTTGDPFGGPPV 159
DB 61 PVIAENAPQPVRIINPVGGFSAFALPAGWVESDAARHFDYGSALLSKTTGDPFGGPPV 120

QY 160 ANDTRIVLGRDQKLYASAEATDSKAAARLGSDMGDFYFYPGTRINQETVSLDANGVSG 219
DB 121 ANDTRIVLGRDQKLYASAEATDSKAAARLGSDMGDFYFYPGTRINQETVSLDANGVSG 180

QY 220 SASYYEVKFSDFPSKNGQIWTGVIGSPAANAPDAGPPQRFVFWLGTANNPVDKGAAL 279
DB 181 SASYYEVKFSDFPSKNGQIWTGVIGSPAANAPDAGPPQRFVFWLGTANNPVDKGAAL 240

QY 280 AESIRPLVAPPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 325
DB 241 AESIRPLVAPPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 286

RESULT 8
US-10-193-002-350
; Sequence 350, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TITLE OF INVENTION: TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/193,002
; FILING DATE: 10-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 350:
; SEQUENCE CHARACTERISTICS:

```

```

;
; LENGTH: 652 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 350:
US-10-193-002-350

Query Match      86.8%; Score 1499; DB 14; Length 652;
Best Local Similarity 100.0%; Pred. No. 1.5e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPAAPADPNAPPVTAAPN 106
DB 374 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPAAPADPNAPPVTAAPN 433
QY 107 APQVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPPVANDTRIV 166
DB 434 APQVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPPVANDTRIV 493
QY 167 LGRLDOKLYASAEATDSKAAARLGSMDGGEFYMPYPGTRINQETVSLDANGVSGSASYEV 226
DB 494 LGRLDOKLYASAEATDSKAAARLGSMDGGEFYMPYPGTRINQETVSLDANGVSGSASYEV 553
QY 227 KFSDSPKNGQIWTGVIQSPAAANADAGPPQRFVWVLGTANNPVVDKGAALAEIRPL 286
DB 554 KFSDSPKNGQIWTGVIQSPAAANADAGPPQRFVWVLGTANNPVVDKGAALAEIRPL 613
QY 287 VAPPAPAPAPAEAPAPAPAGEVAPTTTTPORTLPA 325
DB 614 VAPPAPAPAPAEAPAPAPAGEVAPTTTTPORTLPA 652

RESULT 9
US-10-084-843-355
; Sequence 355, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Applicant: Skeiky, Yasir A.W.
; Dillon, Davin C.
; Campos-Neto, Antonio
; Houghton, Raymond
; Twardzik, Thomas S.
; Lodes, Michael J.
; Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.41109

```

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 652 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 355:
US-10-084-843-355

Query Match      86.8%; Score 1499; DB 14; Length 652;
Best Local Similarity 100.0%; Pred. No. 1.5e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPAAPADPNAPPVTAAPN 106
DB 374 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPAAPADPNAPPVTAAPN 433
QY 107 APQVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPPVANDTRIV 166
DB 434 APQVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPPVANDTRIV 493
QY 167 LGRLDOKLYASAEATDSKAAARLGSMDGGEFYMPYPGTRINQETVSLDANGVSGSASYEV 226
DB 494 LGRLDOKLYASAEATDSKAAARLGSMDGGEFYMPYPGTRINQETVSLDANGVSGSASYEV 553
QY 227 KFSDSPKNGQIWTGVIQSPAAANADAGPPQRFVWVLGTANNPVVDKGAALAEIRPL 286
DB 554 KFSDSPKNGQIWTGVIQSPAAANADAGPPQRFVWVLGTANNPVVDKGAALAEIRPL 613
QY 287 VAPPAPAPAPAEAPAPAPAGEVAPTTTTPORTLPA 325
DB 614 VAPPAPAPAPAEAPAPAPAGEVAPTTTTPORTLPA 652

RESULT 10
US-09-287-849-10
; Sequence 10, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Applicant: Skeiky, Yasir A.W.
; Dillon, Davin C.
; Applicant: Alderson, Mark
; Applicant: Campos-Neto, Antonio
; Applicant: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tetra-fusion
US-09-287-849-10

```

Query Match		86.8%;	Score 1499;	DB 9;	Length 802;	
Best Local Similarity		100.0%;	Pred. No. 1.9e-90;			
Matches 279;		Conservative	0;	Mismatches	0;	Indels
					0;	Gaps
						0;
QY	47	VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAAAPPPADPNAPPPVIAPN	106			
DB	524	VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAAAPPPADPNAPPPVIAPN	583			
QY	107	APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDDPPFCQPPPVANDTRIV	166			
DB	584	APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDDPPFCQPPPVANDTRIV	643			
QY	167	LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV	226			
DB	644	LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV	703			
QY	227	KFSDPSKNGQIWTGIVGSPAANADAGPQBFVWVWLTANNPVDKGAALAESIRPL	286			
DB	704	KFSDPSKNGQIWTGIVGSPAANADAGPQBFVWVWLTANNPVDKGAALAESIRPL	763			
QY	287	VAPPPAPAPAPAPAPAPAGEVAPTPTTPTPQRTPLA	325			
DB	764	VAPPPAPAPAPAPAPAPAGEVAPTPTTPTPQRTPLA	802			

RESULT 11

US-10-193-002-209

Sequence 209, Application US/10193002

Publication No. US20030135026A1

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

Skeiky, Yasir A.W.

Dillon, Davin C.

Campos-Neto, Antonia

Houghton, Raymond

Vedvick, Thomas S.

Twardzik, Daniel R.

Lodes, Michael J.

Hendrickson, Ronald C.

TITL OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS

NUMBER OF SEQUENCES: 350

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/193,002

FILING DATE: 10-Jul-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/072,596

FILING DATE: 05-MAY-1998

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.417C9

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 209:

SEQUENCE CHARACTERISTICS:

LENGTH: 802 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 209:

US-10-193-002-209

Query Match

86.8%;

Score 1499;

DB 14;

Length 802;

Best Local Similarity

100.0%;

Pred. No. 1.9e-90;

Matches 279;

Conservative

0;

Mismatches

0;

Indels

0;

Gaps

0;

QY 47 | VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAAAPPPADPNAPPPVIAPN | 106 || DB | 524 | VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAAAPPPADPNAPPPVIAPN | 583 |
QY	107	APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDDPPFCQPPPVANDTRIV	166
DB	584	APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDDPPFCQPPPVANDTRIV	643
QY	167	LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV	226
DB	644	LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV	703
QY	227	KFSDPSKNGQIWTGIVGSPAANADAGPQBFVWVWLTANNPVDKGAALAESIRPL	286
DB	704	KFSDPSKNGQIWTGIVGSPAANADAGPQBFVWVWLTANNPVDKGAALAESIRPL	763
QY	287	VAPPPAPAPAPAPAPAPAGEVAPTPTTPTPQRTPLA	325
DB	764	VAPPPAPAPAPAPAPAPAGEVAPTPTTPTPQRTPLA	802

```
; INFORMATION FOR SEQ ID NO: 346:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 802 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 346:
US-10-193-002-346

Query Match      86.8%; Score 1499; DB 14; Length 802;
Best Local Similarity 100.0%; Pred. No. 1.9e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPNAAAPPADPNAPPPVIAPN 106
Db 524 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPNAAAPPADPNAPPPVIAPN 583
QY 107 APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPPVANDTRIV 166
Db 584 APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPPVANDTRIV 643
QY 167 LGRLDQKLYASABATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV 226
Db 644 LGRLDQKLYASABATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV 703
QY 227 KFSDPKSKNGQIWTGVI GSPAAAPDAGPPQRFVFWLGTANNPVVDKGAALAEIRPL 286
Db 704 KFSDPKSKNGQIWTGVI GSPAAAPDAGPPQRFVFWLGTANNPVVDKGAALAEIRPL 286
QY 287 VAPPPAPAPAPAPAPAPAPAGEVAPTTPTTPTQRTLPA 325
Db 764 VAPPPAPAPAPAPAPAPAPAGEVAPTTPTTPTQRTLPA 802

RESULT 13
US-10-084-843-214
; Sequence 214, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Skeiky, Yasir A.W.
; Dillon, Davin C.
; Campos-Neto, Antonio
; Houghton, Raymond
; Vedvick, Thomas S.
; Twardzik, Daniel R.
; Lodes, Michael J.
; Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.

; INFORMATION FOR SEQ ID NO: 346:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 802 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 346:
US-10-193-002-346

Query Match      86.8%; Score 1499; DB 14; Length 802;
Best Local Similarity 100.0%; Pred. No. 1.9e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPNAAAPPADPNAPPPVIAPN 106
Db 524 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPNAAAPPADPNAPPPVIAPN 583
QY 107 APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPPVANDTRIV 166
Db 584 APQPVRIIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPQPPPVANDTRIV 643
QY 167 LGRLDQKLYASABATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV 226
Db 644 LGRLDQKLYASABATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV 703
QY 227 KFSDPKSKNGQIWTGVI GSPAAAPDAGPPQRFVFWLGTANNPVVDKGAALAEIRPL 286
Db 704 KFSDPKSKNGQIWTGVI GSPAAAPDAGPPQRFVFWLGTANNPVVDKGAALAEIRPL 286
QY 287 VAPPPAPAPAPAPAPAPAPAGEVAPTTPTTPTQRTLPA 325
Db 764 VAPPPAPAPAPAPAPAPAPAGEVAPTTPTTPTQRTLPA 802

RESULT 14
US-10-084-843-351
; Sequence 351, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Skeiky, Yasir A.W.
; Dillon, Davin C.
; Campos-Neto, Antonio
; Houghton, Raymond
; Vedvick, Thomas S.
; Twardzik, Daniel R.
; Lodes, Michael J.
; Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
```

```
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
;   NAME: Maki, David J.
;   REGISTRATION NUMBER: 31,392
;   REFERENCE/DOCKET NUMBER: 210121.411C9
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (206) 622-4900
;   TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 351:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 802 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 351:
US-10-084-843-351

Query Match      86.8%; Score 1499; DB 14; Length 802;
Best Local Similarity 100.0%; Pred. No. 1.9e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAPADPNAPPVIAPN 106
Db 524 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAPADPNAPPVIAPN 583
QY 107 APOQVRIDNPVGGSFALPAGWVESDAAHFDYGSALLSKTTGDPFPFQPPVANDTRIV 166
Db 584 APOQVRIDNPVGGSFALPAGWVESDAAHFDYGSALLSKTTGDPFPFQPPVANDTRIV 643
QY 167 LGRLDQKLYASAEATDSKAAARLGSDMGCFYMPYPGTRINQETVSLDANGVSGSASYEV 226
Db 644 LGRLDQKLYASAEATDSKAAARLGSDMGCFYMPYPGTRINQETVSLDANGVSGSASYEV 703
QY 227 KFSDDPSKPNQGIWTGIVGSPAANADPAGPPQRFVFWLGTANNPVDKGAALAESIRPL 286
Db 704 KFSDDPSKPNQGIWTGIVGSPAANADPAGPPQRFVFWLGTANNPVDKGAALAESIRPL 763
QY 287 VAPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTLPA 325
Db 764 VAPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTLPA 802

Search completed: July 7, 2004, 18:33:33
Job time : 44.5 secs

; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
;   NAME: Maki, David J.
;   REGISTRATION NUMBER: 31,392
;   REFERENCE/DOCKET NUMBER: 210121.411C9
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (206) 622-4900
;   TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 351:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 802 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 351:
US-10-084-843-351

Query Match      86.8%; Score 1499; DB 14; Length 802;
Best Local Similarity 100.0%; Pred. No. 1.9e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAPADPNAPPVIAPN 106
Db 524 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAPADPNAPPVIAPN 583
QY 107 APOQVRIDNPVGGSFALPAGWVESDAAHFDYGSALLSKTTGDPFPFQPPVANDTRIV 166
Db 584 APOQVRIDNPVGGSFALPAGWVESDAAHFDYGSALLSKTTGDPFPFQPPVANDTRIV 643
QY 167 LGRLDQKLYASAEATDSKAAARLGSDMGCFYMPYPGTRINQETVSLDANGVSGSASYEV 226
Db 644 LGRLDQKLYASAEATDSKAAARLGSDMGCFYMPYPGTRINQETVSLDANGVSGSASYEV 703
QY 227 KFSDDPSKPNQGIWTGIVGSPAANADPAGPPQRFVFWLGTANNPVDKGAALAESIRPL 286
Db 704 KFSDDPSKPNQGIWTGIVGSPAANADPAGPPQRFVFWLGTANNPVDKGAALAESIRPL 763
QY 287 VAPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTLPA 325
Db 764 VAPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTLPA 802

RESULT 15
US-10-359-460-10
; Sequence 10, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
;   APPLICANT: Reed, Steven G.
;   APPLICANT: Skeiky, Yasir A.W.
;   APPLICANT: Dillon, Davin C.
;   APPLICANT: Alderson, Mark
;   APPLICANT: Campos-Neto, Antonio
;   APPLICANT: Corixa Corporation
;   TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
;   TITLE OF INVENTION: and Their Uses
;   FILE REFERENCE: 014058-009020US
;   CURRENT APPLICATION NUMBER: US/10/359,460
;   CURRENT FILING DATE: 2003-02-05
;   PRIOR APPLICATION NUMBER: US/09/287,849
;   PRIOR FILING DATE: 1999-04-07
;   PRIOR APPLICATION NUMBER: US 08/818,112
;   PRIOR FILING DATE: 1997-03-13
;   PRIOR APPLICATION NUMBER: US 08/942,578
;   PRIOR FILING DATE: 1997-10-01
;   PRIOR APPLICATION NUMBER: US 09/025,197
;   PRIOR FILING DATE: 1998-02-18
;   PRIOR APPLICATION NUMBER: US 09/056,556
;   PRIOR FILING DATE: 1998-04-07
;   PRIOR APPLICATION NUMBER: US 09/223,040
;   PRIOR FILING DATE: 1998-12-30
;   NUMBER OF SEQ ID NOS: 46
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 18:17:14 ; Search time 50.5 seconds
(without alignments)
1818.375 Million cell updates/sec

Title: us-10-720-192-3

Perfect score: 1726

Sequence: 1 MHQVDPNLTFRKGRILAALAI.....PAGEVAPPTPTPTPORTLPA 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq 29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1726	100.0	325	2 AAW03565	Aaw03565 Mycobacte
2	1726	100.0	325	5 AAM50732	Aam50732 Mycobacte
3	1726	100.0	332	2 AAW32350	Aaw32350 Mycobacte
4	1726	100.0	332	2 AAW32418	Aaw32418 Mycobacte
5	1726	100.0	332	2 AAW64322	Aaw64322 Mycobacte
6	1726	100.0	332	2 AAW81683	Aaw81683 M. tuberc
7	1726	100.0	332	2 AAY38945	Aay38945 M. tuberc
8	1726	100.0	332	2 AAY39083	Aay39083 M. tuberc
9	1726	100.0	332	2 AAU01895	Aau01895 M. tuberc
10	1726	100.0	332	5 AAE29720	Aae29720 Mycobacte
11	1726	100.0	332	5 AAE17584	Aae17584 Mycobacte
12	1720	99.7	325	3 AAY97281	Aay97281 Fibronect
13	1542	89.3	286	2 AAW03566	Aaw03566 Mycobacte
14	1499	86.8	652	2 AAY39082	Aay39082 M. tubercu
15	1499	86.8	652	2 AAY39225	Aay39225 M. tuberc
16	1499	86.8	802	2 AAW64379	Aaw64379 Mycobacte
17	1499	86.8	802	2 AAW81746	Aaw81746 M. tuberc
18	1499	86.8	802	2 AAY32063	Aay32063 Mycobacte
19	1499	86.8	802	2 AAY39033	Aay39033 M. tuberc
20	1499	86.8	802	2 AAY39081	Aay39081 M. tubercu
21	1499	86.8	802	2 AAY39176	Aay39176 M. tuberc
22	1499	86.8	802	2 AAY39224	Aay39224 M. tuberc
23	1499	86.8	802	5 AAU74592	Aau74592 Antigenic
24	980	56.8	287	7 ADB74347	Adb74347 Mycobacte
25	237.5	13.8	268	4 ABUS2942	Abu52942 Human mam

26	235	13.6	256	4 ABUS2945	Abu52945 Human mam
27	233.5	13.5	262	4 ABUS2944	Abu52944 Human mam
28	233	13.5	270	4 ABUS2939	Abu52939 Human mam
29	233	13.5	8991	6 ABUS2939	Abu52939 Human mam
30	228	13.2	544	3 AAG15453	Aag15453 Arabidops
31	227.5	13.2	652	7 ADB65210	Adb65210 Human pro
32	227	13.2	598	6 ABUS37030	Abu37030 Protein e
33	225.5	13.1	446	4 ABB70063	Abb70063 Drosophil
34	223.5	12.9	652	5 ABB92424	Abb92424 Herbigida
35	222.5	12.9	267	4 ABUS2938	Abu52938 Human mam
36	222	12.9	763	2 AAW31852	Aaw31852 Mycobacte
37	221.5	12.8	260	4 ABUS2932	Abu52932 Human mam
38	221.5	12.8	267	4 ABUS2931	Abu52931 Human mam
39	218.5	12.7	253	4 ABUS2937	Abu52937 Human mam
40	217.5	12.6	817	6 ABR53281	Abr53281 Protein s
41	217	12.6	261	4 ABUS2940	Abu52940 Human mam
42	217	12.6	277	4 ABUS2936	Abu52936 Human mam
43	216.5	12.5	254	4 ABUS2933	Abu52933 Human mam
44	214	12.4	406	4 ABG27250	Abg27250 Novel hum
45	212.5	12.3	258	4 ABUS2941	Abu52941 Human mam

ALIGNMENTS

RESULT 1
AAW03565
ID AAW03565 standard; protein; 325 AA.
XX
AC AAW03565;
XX
DT 22-APR-1997 (first entry)
XX
DE Mycobacterium tuberculosis antigenic determinant protein.

XX Antigenic determinant; Mycobacterium tuberculosis; ion exchange; human;
KW chromatography; gel filtration; reverse phase column chromatography;
KW immunogenic; serum; guinea pig; expression vector; cosmid; antibody;
KW Mycobacterium smegmatis; Mycobacterium bovis BCG; microorganism; vaccine;
KW hybrid; epitope; disease; diphtheria; cholera; toxin.
XX Mycobacterium tuberculosis.
OS Mycobacterium tuberculosis.

FH Key Location/Qualifiers
FT Peptide 1..39
FT /note= "signal peptide"
FT Protein 40..325
FT /note= "mature protein"

WO9623985-A1.

08-AUG-1996.

31-JAN-1996; 96WO-FR000166.

01-FEB-1995; 95US-00362184.

(INSP) INST PASTEUR.

Laqueyrie A, Marchal G, Pescher P, Romain F;

WPI; 1996-371433/37.

N-PSDB; AAT39357.

Mycobacterium tuberculosis antigen and hybrid proteins comprising it -
useful in vaccines against tuberculosis and in immunoassays.

Claim 2; Page 49-50; 74pp; French.

This is the amino acid sequence of an antigenic determinant protein from
Mycobacterium tuberculosis. The mature protein has calculated mol. wt. of
28779 Da but has an experimental mol. wt of 45-47 kD as determined by SDS
-PAGE. The difference is thought to be due to the high frequency of Pro

residues (21.7%) in the sequence. The protein was purified from *M. tuberculosis* strain H37Rv by conventional chromatographic methods e.g. low pressure ion exchange chromatography, Si 300 gel filtration, DEAE ion exchange and reverse phase column chromatography. The proteins were then separated by SDS-PAGE and immunogenic proteins determined with immune serum from humans with TB or guinea pigs infected with *M. tuberculosis*. A complex of mol. wt. 45-47 kD was identified. An expression library of genomic *M. tuberculosis* DNA was generated in the cosmid vector pUml8. The library was transformed into *M. smegmatis*. Clones expressing antigenic determinants were isolated by screening with an antibody against the *M. bovis* BCG 45/47 kD proteins. 3 clones were isolated, all of which contained identical sequences i.e. the sequence presented here. The protein or microorganisms expressing it can be used as vaccines against tuberculosis. Also hybrid proteins comprising this protein and epitopes from other disease causing organisms or proteins, e.g. diphtheria or cholera toxin, can be used as vaccines against their respective diseases

Query Match	100.0%;	Score	1726;	DB	2;	Length	325;		
Best Local Similarity	100.0%;	Pred. No.	1.3e-104;						
Matches	325;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MHOVDPNITREKGLAALA	TAAMASASLVTVAVPATANADPEAPVPVTTAASPSPTAAA	60					
DB	1	MHOVDPNLITRRKGLAALA	TAAMASASLVTVAVPATANADPEAPVPVTTAASPSPTAAA	60					
QY	61	PPAPATPVAPPPPPAAAAANTNAQPGDNAAAPPADPNAPPPVIAENAPQPVRI	DNPVGGF	120					
DB	61	PPAPATPVAPPPPPPPAAANTNAQPGDNAAAPPADPNAPPPVIAENAPQPVRI	DNPVGGF	120					
QY	121	SFALPAGWVSDDAAHFDYGSALLSKTGTDPDPFGGPPVPVANDTRIVLGR	LQKUYASAEA	180					
DB	121	SFALPAGWVSDDAAHFDYGSALLSKTGTDPDPFGGPPVPVANDTRIVLGR	LQKUYASAEA	180					
QY	181	TDSKAAARLSDMGGEFVMPYFGTRINQETVSLDANGVSGSASYEVKESD	SKNGQIWT	240					
DB	181	TDSKAAARLSDMGGEFVMPYFGTRINQETVSLDANGVSGSASYEVKESD	SKNGQIWT	240					
QY	241	GVIGSPAANADGAGPPQORFVFWLGTANNPVDKGAAKALABSIRPLVAP	PPPPAPAPAPAE	300					
DB	241	GVIGSPAANADGAGPQORFVFWLGTANNPVDKGAAKALABSIRPLVAP	PPPPAPAPAPAE	300					
QY	301	APAPAPAGEVAPTPTTPTBORTLPA	325						
DB	301	APAPAPAGEVAPTPTTPTBORTLPA	325						

RESULT 2	
AAW50732	
AAW50732	standard; protein; 325 AA.
XX	
XX	AAW50732;
XX	
XX	18-APR-2002 (first entry)
XX	
XX	Mycobacterium tuberculosis immunodominant Mtb protein ModD MPT32.
XX	
XX	Mtb; ModD; MPT32; immunogen; mycobacteria; immunisation; vaccine.
XX	
XX	Mycobacterium tuberculosis.
XX	
XX	Key
XX	Location/Qualifiers
XX	1. 39
XX	/label= Signal peptide
XX	/note= "SPC-dependent signal secretion sequence"
XX	40. 325
XX	/label= Mature_protein
XX	
XX	WO200204018-A2.
XX	
XX	17-JAN-2002.
XX	

10-JUL-2001; 2001WO-US021717.
10-JUL-2000; 2000US-0217646P.
(COLS) UNIV COLORADO STATE RES FOUND.
Orme IM, Belisle JT;
WPI; 2002-164602/21.
Vaccine for boosting immunity to mycobacteria when administered in mid-life in a subject who has been vaccinated in childhood with Bacillus Calmette-Guerrin, has purified proteins from mycobacterium tuberculosis.
Claim 8; Page 17; 61pp; English.
The present sequence is that of the Mycobacterium tuberculosis (Mtb) strain H37Rv gene Rv1860 product, designated ModD MPT32. This is one of 31 immunodominant secreted or cytosolic Mtb proteins of strain H37Rv (see AAM50729-59) discovered through the use of 2-dimensional liquid phase electrophoresis coupled with an in vitro interferon-gamma assay and liquid chromatography-mass spectrometry. The immunogens stimulate a strong interferon-gamma response from T cells of M. tuberculosis infected mice. The invention provides vaccine compositions for boosting immunity to mycobacteria when administered in mid-life to a subject who has been vaccinated neonatally or in early childhood with BCG and in whom protective immunity has waned. The vaccine compositions comprise 1 or more of the 31 purified immunogenic proteins. When used as immunogens, the secreted Mtb proteins lack the secreted signal sequence. A preferred protein is Ag85A (see AAM50759), the secreted product of the Rv3084v gene
Sequence 325 AA;

Query Match	100.0%;	Score 1726;	DB 5;	Length 325;
Beat Local Similarity	100.0%;	Pred. No. 1.3e-104;		
Matches 325;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MHVDPNLTREKGRLLAALAIAAMASGLVTVAVPATANADPEAPPPVPTTAAAGPPSTAAA	60	
DB	1	MHVDPNLTREKGRLLAALAIAAMASGLVTVAVPATANADPEAPPPVPTTAAAGPPSTAAA	60	
QY	61	PPAPATVAPPPPPAAANTENACPGDENAPADPNAPPPEVIAAPNAPQPVRIIDNPVGGF	120	
DB	61	PPAPATVAPPPPPAAANTENACPGDENAPADPNAPPPEVIAAPNAPQPVRIIDNPVGGF	120	
QY	121	SFALPAGWESDAAHFDYGSALLSKTTGDDPPPGQPPPVANDTRIVLGRLDQKLYASAEA	180	
DB	121	SFALPAGWESDAAHFDYGSALLSKTTGDDPPPGQPPPVANDTRIVLGRLDQKLYASAEA	180	
QY	181	TDSKAARLGSDMGEEVMPYPGTRINQETVSLDANGVSGSASYEVEKFSDFSPKNGQIWT	240	
DB	181	TDSKAARLGSDMGEEVMPYPGTRINQETVSLDANGVSGSASYEVEKFSDFSPKNGQIWT	240	
QY	241	GVIGSPAANADAGPPQRFVFWLGTANNPVDKGAAKALAESIRPLVAPPPAPAPAPAE	300	
DB	241	GVIGSPAANADAGPPQRFVFWLGTANNPVDKGAAKALAESIRPLVAPPPAPAPAPAE	300	
QY	301	APAPAGVAPPTPTPTPTORTLPA	325	
DB	301	APAPAGVAPPTPTPTPTORTLPA	325	

RESULT 3	
AAW32350	
ID	AAW32350 standard; protein; 332 AA.
XX	
AC	AAW32350;
XX	
DT	13-JAN-1998 (first entry)
XX	
DE	Mycobacterium tuberculosis antigen DPEP.
XX	
XX	Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;

```

XW skin testing; M.tuberculosis.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9709429-A2.
XX
PD 13-MAR-1997.
XX
PF 30-AUG-1996; 96WO-US014675.
XX
PR 01-SEP-1995; 95US-00523435.
XX
PR 22-SEP-1995; 95US-00532136.
XX
PR 22-MAR-1996; 96US-00620280.
XX
PR 05-JUN-1996; 96US-00658800.
XX
PR 12-JUL-1996; 96US-00680573.
XX
PA (CORI-) CORIXA CORP.
XX
XX
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TH, Twardzik DR;
XX
XX WPI; 1997-192904/17.
XX
XX N-PSDB; AAT91423.
XX
XX New immunogenic polypeptide(s) from soluble M. tuberculosis antigens -
PT useful for diagnosis of M. tuberculosis infection.
XX
XX Example 1; Page 92-94; 190pp; English.
XX
XX A new immunogenic polypeptide has been developed comprising an
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
CC variant differing only in conservative substitutions and/or
CC modifications). The present sequence represents a M.tuberculosis antigen,
CC DPEP. The immunogenic polypeptide can be used to diagnose M.tuberculosis
CC infection by forming complexes with specific antibodies in the sample.
CC Fragments of DNA encoding the immunogenic polypeptide can be used as
CC diagnostic primers or probes and agents that bind to the antigen.
CC especially monoclonal antibodies or equivalent polyclonal antibodies, are
CC also used for diagnosis
XX
XX Sequence 332 AA;
XX
XX Query Match 100.0%; Score 1726; DB 2; Length 332;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-104;
XX Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MHQVDPNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEAPVPTTAAASPTAAA 60
DB 8 MHQVDPNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEAPVPTTAAASPTAAA 67
XX
QY 61 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 120
DB 68 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 127
XX
QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQGPVPPVANDTRIVLGRDLQKLYASAEA 180
DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQGPVPPVANDTRIVLGRDLQKLYASAEA 187
XX
QY 181 TDSKAAARLGSDMGGEFFYMPYPGTRINQETVSLDANGVSGSASYEVKFSDFSKNGQIWT 240
DB 188 TDSKAAARLGSDMGGEFFYMPYPGTRINQETVSLDANGVSGSASYEVKFSDFSKNGQIWT 247
XX
QY 241 GVIGSPAANAPDAGPPQRFVWVLGTANNPVDKGAALKALAESIRPLVAPPAPAPAPAE 300
DB 248 GVIGSPAANAPDAGPPQRFVWVLGTANNPVDKGAALKALAESIRPLVAPPAPAPAPAE 307
XX
QY 301 APAPAPAGEVAPTETTTPTORTLPA 325
DB 308 APAPAPAGEVAPTETTTPTORTLPA 332
XX
RESULT 4
AAW32418

```

```

ID AAW32418 standard; protein; 332 AA.
XX
AC AAW32418;
XX
DT 08-JAN-1998 (first entry)
XX
DE Mycobacterium tuberculosis antigen DPEP.
XX
KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW skin testing; M.tuberculosis.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9709428-A2.
XX
XX 13-MAR-1997.
XX
XX 30-AUG-1996; 96WO-US014674.
XX
XX 01-SEP-1995; 95US-00523436.
XX
XX 22-SEP-1995; 95US-00532136.
XX
XX 22-MAR-1996; 96US-00620874.
XX
XX 05-JUN-1996; 96US-00658800.
XX
XX 12-JUL-1996; 96US-00680574.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Skeiky YA, Dillon DC, Campos-Neto A, Houghton R;
XX Vedvick TH, Twardzik DR;
XX
XX WPI; 1997-192903/17.
XX
XX N-PSDB; AAT91486.
XX
XX New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
PT useful in vaccines for prevention or treatment of tuberculosis, also for
XX diagnosis.
XX
XX Example 1; Page 87-89; 168pp; English.
XX
XX A new immunogenic polypeptide has been developed comprising an
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
CC variant differing only in conservative substitutions and/or
CC modifications). The present sequence represents a M.tuberculosis antigen,
CC DPEP. The immunogenic protein, and fusion proteins containing one or more
CC of the proteins or one of the proteins plus ESAT-6, are useful in
CC vaccines, preferably when formulated with a non-specific adjuvant, to
CC induce an immune response against M.tuberculosis (for treatment or
CC prevention)
XX
XX Sequence 332 AA;
XX
XX Query Match 100.0%; Score 1726; DB 2; Length 332;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-104;
XX Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MHQVDPNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEAPVPTTAAASPTAAA 60
DB 8 MHQVDPNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEAPVPTTAAASPTAAA 67
XX
QY 61 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 120
DB 68 PPAPATPVAPPPPPAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 127
XX
QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQGPVPPVANDTRIVLGRDLQKLYASAEA 180
DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQGPVPPVANDTRIVLGRDLQKLYASAEA 187
XX
QY 181 TDSKAAARLGSDMGGEFFYMPYPGTRINQETVSLDANGVSGSASYEVKFSDFSKNGQIWT 240
DB 188 TDSKAAARLGSDMGGEFFYMPYPGTRINQETVSLDANGVSGSASYEVKFSDFSKNGQIWT 247
XX
QY 241 GVIGSPAANAPDAGPPQRFVWVLGTANNPVDKGAALKALAESIRPLVAPPAPAPAPAE 300

```

Db	248	GVIGSPAANADGAGPQRFWFWLGTANNPVVDKGAAKALAESIRPLVAVPPAPAPAPAE	307
QY	301	APAPAPAGEVAPTPTPTPTFORTLPA	325
Db	308	APAPAPAGEVAPTPTPTPTFORTLPA	332
RESULT 5			
AAW64322	AAW64322 standard; protein; 332 AA.		
XX	AC	AC	AAW64322;
XX	17-OCT-2003	(revised)	
DT	09-NOV-1998	(first entry)	
XX	Mycobacterium tuberculosis antigen DPEP.		
DE	Tuberculosis; infection; diagnosis; antigen; DPEP.		
XX	Mycobacterium tuberculosis; strain H37Rv.		
OS	W09816645-A2.		
XX	23-APR-1998.		
XX	07-OCT-1997;	97WO-US018214.	
PF	11-OCT-1996;	96US-00729622.	
XX	13-MAR-1997;	97US-00818111.	
XX	(CORI-) CORIXA CORP.		
PA			
XX	Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;		
PI	Vedvick TS, Twardzik DR, Lodes MJ;		
XX	WPI; 1998-251232/22.		
DR	N-PSDB; AAW44390.		
XX			
DR	New isolated Mycobacterium tuberculosis polypeptides and DNA - used to		
PT	develop products for the detection of M. tuberculosis infection and		
PT	diagnosis of tuberculosis.		
XX	Example 1; Page 93-94; 250pp; English.		
PS			
XX	This polypeptide comprises the Mycobacterium tuberculosis antigen DPEP. A		
CC	DNA sequence (see AAW44390) coding for DPEP was isolated from a M.		
CC	tuberculosis genomic library using a probe based on an isolated N-		
CC	terminal peptide (see AAW64347). The invention relates to compositions		
CC	and methods for diagnosing tuberculosis. It provides polypeptides (see		
CC	AAW4291-W64379) comprising an antigenic portion of a soluble M.		
CC	tuberculosis antigen, or an immunogenic portion of an M. tuberculosis		
CC	antigen, as well as DNA sequences encoding such polypeptides, recombinant		
CC	expression vectors and transformed or transfected host cells. Also		
CC	claimed are methods and diagnostic kits for detecting M. tuberculosis		
CC	infection in a patient using these polypeptides, antibodies or		
CC	oligonucleotide probes and primers, for the diagnosis of tuberculosis.		
CC	(Updated on 17-OCT-2003 to standardise OS field)		
XX	Sequence 332 AA;		
SQ			
Query Match 100.0%; Score 1726; DB 2; Length 332;			
Best Local Similarity 100.0%; Pred. No. 1.3e-104;			
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	MHQVDPNLTRKGLAALATAAASASLTVAVPATANADPPAPVPVTTAASPESTAAA	60
Db	8	MHQVDPNLTRKGLAALATAAASASLTVAVPATANADPPAPVPVTTAASPESTAAA	67
QY	61	PPAPATVPAPPPPPAAAANTPNAQPGDNAAAPPADPNAPPPVPIANAPQPVRIIDNPVGGF	120
Db	68	PPAPATVPAPPPPPAAAANTPNAQPGDNAAAPPADPNAPPPVPIANAPQPVRIIDNPVGGF	127

D b	68	FPAPATPVAPPPPPAAANTFNAQGDGDNAAAPPADPNAPPVPTAPNAPQFVDRIDNPVG	127
Q y	121	SPALPAGWESDAAHPDYGSALLSKTTGDPFPFGQPPPVANTRIVLGRLDOKLYASAEA	180
D b	128	SPALPAGWESDAAHPDYGSALLSKTTGDPFPFGQPPPVANTRIVLGRLDOKLYASAEA	187
Q y	181	TDSKAAARLGSOWGEFYMYPGTRINQETVSLDANGVSGSASYYEVKESDPKPKNGQIWT	240
D b	188	TDSKAAARLGSOWGEFYMYPGTRINQETVSLDANGVSGSASYYEVKESDPKPKNGQIWT	247
Q y	241	GVIGSPANAPDAGPPQRFVWVLGTANNPVDKGAAKALAESIRPLVAPPPAPAPAPAP	300
D b	248	GVIGSPANAPDAGPPQRFVWVLGTANNPVDKGAAKALAESIRPLVAPPPAPAPAPAP	307
Q y	301	APAPAPAGEVAPTTPTPTPQRTLPA	325
D b	308	APAPAPAGEVAPTTPTPTPQRTLPA	332

RESULT 7
AAY38945
ID AAY38945 standard; protein; 332 AA.

AC	AAV38945;
XX	
DT	05-NOV-1999 (first entry)
XX	
DE	M. tuberculosis recombinant antigen protein pREP.
XX	
KW	Antigen; diagnosis; detection; infection; antibody; immunisation;
KW	vaccine; immunity.
XX	

XX
OS Mycobacterium tuberculosis.
XX
PN WO9942118-A2.

AA	
PD	26-AUG-1999.
XX	
PF	17-FEB-1999; 99WO-US003265.

XX
PR 18-FEB-1998; 98US-00024753.
PR 05-MAY-1998; 98US-00072596.
XX

PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS Twardzik no Pedro m
PI

XX	
DR	WPI; 1999-527416/44.
DR	N-PSDB; AAZ19088.
yy	

PT New polypeptide comprising antigenic portions of M. tuberculosis.
XX
PS Example 1; Page 131-132; 323pp; English.
PV

This invention describes novel recombinant antigens and their encoded nucleic acids derived from *Mycobacterium tuberculosis*. The novel polypeptides are useful for detecting *M. tuberculosis* infection in

polypeptides, and are useful as vaccines for immunizing against M. tuberculosis infection. The new detection methods are needed as current vaccination strategies do not provide 100% immunity

SQ Sequence 332 AA;
Query Match 100.0%; Score 1726; DB 2; Length 332;

Qy 1 MHQVDNLTTRKGRLLAALIAAVASASLTVVATANADPEPAPPVPTTAASPBST
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0; GC Content 50.0%; Freq. NO. 1.3E-104;

D**b** 8 MHQVDNLTRRKGRLAALAIAMASASLVTVAVPATANADPEAPPVPITTAASPEST

Qy	61	PPAPATPVAPPAAAAANTNAQFGDPNAAPPADPNAPPPVIAFNAPQVRI	120
Db	68	PPAPATPVAPPAAAAANTNAQFGDPNAAPPADPNAPPPVIAFNAPQVRI	127
Qy	121	SFALPAGWYESDAAHFDYGSALLSKTTGDPPFPQPPVANDTRIVLGRDOKL	180
Db	128	SFALPAGWYESDAAHFDYGSALLSKTTGDPPFPQPPVANDTRIVLGRDOKL	187
Qy	181	TDSKAAARLGSMDGFFYMPYQPTRINQETVSLDANGVSGSASYEYKTS	240
Db	188	TDSKAAARLGSMDGFFYMPYQPTRINQETVSLDANGVSGSASYEYKTS	247
Qy	241	GVIGSPAANAPDAGPQKWFVVLGTANNPVDKGAAKALAESIRPLVAPPP	300
Db	248	GVIGSPAANAPDAGPQKWFVVLGTANNPVDKGAAKALAESIRPLVAPPP	307
Qy	301	APAPAPAGEVAPPTPTTPTFORTLPA	325
Db	308	APAPAPAGEVAPPTPTTPTFORTLPA	332

RESULT 8
AAY39083

ID AAY39083 standard; protein; 332 AA.

AC AAY39083:

XX
DT 05-NOV-1999 (first entry)

DE M. tuberculosis antigen DPEP amino acid sequence.

KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW immunotherapy; diagnosis; immunisation; vaccine; infection;
KW immune response; skin test.

OS *Mycobacterium tuberculosis*.

XX
PN WO9942076-A2.

XX
PD
26-AUG-1999.

XX
PF 17-FEB-1999; 99WO-US003268.

PR 18-FEB-1998; 98US-00025197.

PR 05-MAY-1998; 98US-00072967.
XX

PA (CORI-) CORIXA CORP.

PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;

DR WPI; 1999-527409/44.
DR N-PSDE; AAZ19300.

PT New antigens from *Mycobacterium tuberculosis* useful in diagnostic skin tests and protective or therapeutic vaccines or compositions.

PS Example 1; Page 93; 299pp; English.

The present invention describes polypeptides comprising an immunogenic part of a Mycobacterium tuberculosis antigen (Ag). Also described are vaccines and fusion protein containing M. tuberculosis Ag's. M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other polypeptides fragments, can be used in pharmaceutical compositions or vaccines to generate a protective or therapeutic immune response to M. tuberculosis and as reagents in skin tests for diagnosis of tuberculosis. Ag can induce proliferation of, or cytokine secretion by, T_H or natural killer cells and/or macrophages in tuberculosis-immune subjects. AA19249 to AA213460 and AA239083 to AA239225 are used in the exemplification of the present invention.

Sequence 332 AA:

55

Query Match 100.0%; Score 1726; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 1.3e-104;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDNLTERRKRLAALAIAAMASASLVTVAVPATANADPPAPPVPTTAASPPSTAAA 60
 DB 8 MHQVDNLTERRKRLAALAIAAMASASLVTVAVPATANADPPAPPVPTTAASPPSTAAA 67
 QY 61 PPAPATVAPPPPPAAANTPNAQPGDNNAPPPADPNAPPVPIAPNAPQVRIIDNPVGGF 120
 DB 68 PPAPATVAPPPPPAAANTPNAQPGDNNAPPPADPNAPPVPIAPNAPQVRIIDNPVGGF 127
 QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFGQPPPVANDTRIVLGRDQKLYASAEA 180
 DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFGQPPPVANDTRIVLGRDQKLYASAEA 187
 QY 181 TDSKAAARLGSDMGCFYMPYPTGTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 240
 DB 188 TDSKAAARLGSDMGCFYMPYPTGTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 247
 QY 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAEIRPLVAPPAPAPAPAP 300
 DB 248 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAEIRPLVAPPAPAPAPAP 307
 QY 301 APAPAPAGEVAPTPTTPTPQRTPLA 325
 DB 308 APAPAPAGEVAPTPTTPTPQRTPLA 332

RESULT 9
 AAU01895
 ID AAU01895 standard; protein; 332 AA.
 XX
 AC AAU01895;
 XX
 DT 29-AUG-2001 (first entry)
 XX
 DE M. tuberculosis DPEP antigen.
 XX
 DE DPEP; antigen; vaccine; tuberculosis; AIDS;
 XX
 KW acquired immunodeficiency disease.
 XX
 OS Mycobacterium tuberculosis.
 XX
 FN WO200124820-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 10-OCT-2000; 2000WO-US028095.
 XX
 PR 07-OCT-1999; 99US-0158338P.
 XX
 PR 07-OCT-1999; 99US-0158425P.
 XX
 FA (CORI-) CORIXA CORP.
 XX
 XX Skeiky Y, Reed S, Houghton RL, McNeill PD, Dillon DC, Lodes ML;
 XX
 XX WPI: 2001-290576/30.
 XX
 XX N-PSDB; AAS03786.
 XX
 XX Vaccinating against Mycobacteria infections in mammals using fusion
 XX
 XX proteins comprising combinations of heterologous antigens.
 XX
 XX PS Disclosure; Page 160; 168pp; English.
 XX
 XX The sequence represents Mycobacterium tuberculosis DPEP, an M.
 XX
 XX tuberculosis antigen. Compositions comprising at least 2 heterologous
 XX
 XX antigens, as a fusion protein, and vectors expressing the fusion proteins
 XX
 XX are used as vaccines to prophylactically immunise mammals (especially
 XX
 XX humans) against infection by Mycobacteria. The compositions contain at
 XX
 XX least 2 heterologous antigens that increase the serological sensitivity
 XX
 XX of individuals infected with tuberculosis, a disease frequently affecting

CC patients with acquired immunodeficiency disease, AIDS
 XX
 SQ Sequence 332 AA;

Query Match 100.0%; Score 1726; DB 4; Length 332;
 Best Local Similarity 100.0%; Pred. No. 1.3e-104;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDNLTERRKRLAALAIAAMASASLVTVAVPATANADPPAPPVPTTAASPPSTAAA 60
 DB 8 MHQVDNLTERRKRLAALAIAAMASASLVTVAVPATANADPPAPPVPTTAASPPSTAAA 67
 QY 61 PPAPATVAPPPPPAAANTPNAQPGDNNAPPPADPNAPPVPIAPNAPQVRIIDNPVGGF 120
 DB 68 PPAPATVAPPPPPAAANTPNAQPGDNNAPPPADPNAPPVPIAPNAPQVRIIDNPVGGF 127
 QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFGQPPPVANDTRIVLGRDQKLYASAEA 180
 DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFGQPPPVANDTRIVLGRDQKLYASAEA 187
 QY 181 TDSKAAARLGSDMGCFYMPYPTGTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 240
 DB 188 TDSKAAARLGSDMGCFYMPYPTGTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 247
 QY 241 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAEIRPLVAPPAPAPAPAP 300
 DB 248 GVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAALAEIRPLVAPPAPAPAPAP 307
 QY 301 APAPAPAGEVAPTPTTPTPQRTPLA 325
 DB 308 APAPAPAGEVAPTPTTPTPQRTPLA 332

RESULT 10
 AAE29720
 ID AAE29720 standard; protein; 332 AA.
 XX
 AC AAE29720;
 XX
 DT 27-JAN-2003 (first entry)
 XX
 DE Mycobacterium tuberculosis DPEP antigenic protein.
 XX
 XX Vaccine; immunity; diagnostic agent; gene therapy; DPEP antigen.
 XX
 OS Mycobacterium tuberculosis.
 XX
 FN WO200272792-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 13-MAR-2002; 2002WO-US008223.
 XX
 PR 13-MAR-2001; 2001US-0275837P.
 XX
 FA (CORI-) CORIXA CORP.
 XX
 XX Skeiky Y, Brannon M, Guderian J;
 XX
 XX WPI: 2002-759844/82.
 XX
 XX N-PSDB; AAD47097.
 XX
 XX New recombinant nucleic acid molecule comprising a Leishmania TSA, LeifF,
 XX
 XX M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
 XX
 XX against pathogenic microorganisms e.g. Leishmania and Mycobacterium
 XX
 XX tuberculosis.
 XX
 XX Disclosure; Page 111; 155pp; English.
 XX
 XX The invention relates to a recombinant nucleic acid molecule encoding a
 XX
 XX fusion polypeptide. The recombinant nucleic acid comprises a heterologous
 XX
 XX polynucleotide sequence encoding an antigen or an antigenic fragment from
 XX
 XX Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a

CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, Leif, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides, as
CC in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is M.
CC tuberculosis DPEP antigenic protein
XX
SQ Sequence 332 AA;

Query Match 100.0%; Score 1726; DB 5; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.3e-104;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEAPVPVTTAASPSTAAA 60
Db 8 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEAPVPVTTAASPSTAAA 67
QY 61 PPAPATPVAPPPAAANTPNAQPGDPAAPPPADPNAPPPVVIAPNAPQPVRIIDNPVGGF 120
Db 68 PPAPATPVAPPPAAANTPNAQPGDPAAPPPADPNAPPPVVIAPNAPQPVRIIDNPVGGF 127
QY 121 SPALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPVANDTRIVLGRDOKLYASAEA 180
Db 128 SPALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPVANDTRIVLGRDOKLYASAEA 187
QY 181 TDSKAAARLGSDMGDFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDPKNGQIWT 240
Db 188 TDSKAAARLGSDMGDFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDPKNGQIWT 247
QY 241 GVIGSPAANADPAGPPQRFVWVLTANNPVDKGAALAESIRPLVAPPPAPAPAPAP 300
Db 248 GVIGSPAANADPAGPPQRFVWVLTANNPVDKGAALAESIRPLVAPPPAPAPAPAP 307
QY 301 APAPAPAGEVAPTPTTPTPTQRTLPA 325
Db 308 APAPAPAGEVAPTPTTPTPTQRTLPA 332

RESULT 11
AAE17584
ID AAE17584 standard; protein; 332 AA.
XX
AC AAE17584;
XX
DT 22-APR-2002 (first entry)
XX
DE Mycobacterium species DPEP protein.
XX
KW Fusion protein; antigen; serological sensitivity; immune response;
KW tuberculosis; infection; vaccine; DPEP protein.
XX
OS Mycobacterium sp.
XX
FN WO200198460-A2.
XX
PD 27-DEC-2001.
XX
PF 20-JUN-2001; 2001WO-US019959.
XX
PR 20-JUN-2000; 2000US-00597796.
PR 01-FEB-2001; 2001US-0265737P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Reed S, Alderson M;
XX
DR WPI; 2002-147798/19.
DR N-PSDB; AAD28355.
XX
PT Composition comprising MTB39 antigen and MTB32A antigen from

PT Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS Claim 9; Page 127; 136pp; English.
XX
CC The present invention relates to fusion proteins containing at least two
CC Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human.
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected with
CC Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is Mycobacterium species
CC DPEP protein
XX
SQ Sequence 332 AA;

Query Match 100.0%; Score 1726; DB 5; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.3e-104;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEAPVPVTTAASPSTAAA 60
Db 8 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEAPVPVTTAASPSTAAA 67
QY 61 PPAPATPVAPPPAAANTPNAQPGDPAAPPPADPNAPPPVVIAPNAPQPVRIIDNPVGGF 120
Db 68 PPAPATPVAPPPAAANTPNAQPGDPAAPPPADPNAPPPVVIAPNAPQPVRIIDNPVGGF 127
QY 121 SPALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPVANDTRIVLGRDOKLYASAEA 180
Db 128 SPALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPVANDTRIVLGRDOKLYASAEA 187
QY 181 TDSKAAARLGSDMGDFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDPKNGQIWT 240
Db 188 TDSKAAARLGSDMGDFYMPYPTGTRINQETVSLDANGVSGSASYEVKFSDPKNGQIWT 247
QY 241 GVIGSPAANADPAGPPQRFVWVLTANNPVDKGAALAESIRPLVAPPPAPAPAPAP 300
Db 248 GVIGSPAANADPAGPPQRFVWVLTANNPVDKGAALAESIRPLVAPPPAPAPAPAP 307
QY 301 APAPAPAGEVAPTPTTPTPTQRTLPA 325
Db 308 APAPAPAGEVAPTPTTPTPTQRTLPA 332

RESULT 12
AAE17584
ID AAE17584 standard; protein; 325 AA.
XX
AC AAE17584;
XX
DT 03-JAN-2001 (first entry)
XX
DE Fibronectin attachment protein FAP-B.
XX
KW Inflammation; inflammatory response; irritant; pathogen; treatment;
KW T helper cell; lymphocyte; cell mediated immunity; skin allergy; hives;
KW allergic rhinitis; conjunctivitis; hay fever; allergic gastroenteritis;
KW asthma; bronchopulmonary aspergillosis; pollutant;
KW respiratory tract infection.
XX

Qy	227	KFSDPSKNGQIWTGVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAAKALAESIRPL	286
Db	554	KFSDPSKNGQIWTGVIGSPAANAPDAGPPQRFVWVWLTANNPVDKGAAKALAESIRPL	613
Qy	287	VAPPPAPAPAPAPAPAPAGEVAPTPTTPTQRTLPA	325
Db	614	VAPPPAPAPAPAPAPAPAGEVAPTPTTPTQRTLPA	652

Search completed: July 7, 2004, 18:21:03
 Job time : 50.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 18:19:15 ; Search time 22.5 Seconds
(without alignments)
1389.433 Million cell updates/sec

Title: US-10-720-192-3
Perfect score: 1726
Sequence: 1 MHQVDPNLTTRKGRLLAALAI.....PAGEVAPPTPTPTPQRTLPA 325

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1726	100.0	325	2 D70666	probable modD prot
2	980	56.8	287	2 B87166	hypothetical prote
3	241	14.0	1585	2 T31611	hypothetical prote
4	237.5	13.8	1188	2 S49915	extensin-like prot
5	228.5	13.2	464	2 S22697	extensin - Volvox
6	226.5	13.1	222	2 H96711	hypothetical prote
7	224.5	13.0	3534	2 T42587	tegument protein 2
8	219.5	12.7	85	2 T44744	antigen 43L [impor
9	217.5	12.6	817	2 S51342	verprolin - yeast
10	217.5	12.6	839	2 T75518	hypothetical prote
11	217	12.5	801	2 T29018	hypothetical prote
12	216.5	12.5	847	2 P96531	hypothetical prote
13	213.5	12.4	760	2 T06291	extensin homolog 1
14	213.5	12.4	3164	1 WNB832	U136 protein - hum
15	208.5	12.1	214	2 T10737	extensin-like cell
16	207	12.0	416	1 SKXLAG	dermal gland prote
17	206.5	12.0	1611	2 T38236	hypothetical prote
18	204.5	11.8	539	2 T28770	hypothetical prote
19	203	11.8	1201	2 G86441	unknown protein [i
20	203	11.8	3421	1 W2B836	367K tegument prot
21	202.5	11.7	283	2 S13383	hydroxyproline-ric
22	201.5	11.7	395	2 H75457	hypothetical prote
23	200.5	11.6	214	2 T09854	proline-rich cell
24	200.5	11.6	760	2 T08637	probable Pto kinas
25	199	11.5	708	2 D96711	hypothetical prote
26	199	11.5	839	2 T04859	extensin homolog F
27	197.5	11.4	865	2 A47282	calcium-binding pr
28	196.5	11.4	418	2 T15442	hypothetical prote
29	196.5	11.4	666	2 B70803	hypothetical prote

30	194	11.2	191	2 F84522	probable proline-r
31	193.5	11.2	1151	2 T18535	high molecular mas
32	193	11.2	620	2 S06733	hydroxyproline-ric
33	192.5	11.2	731	2 T04455	hypothetical prote
34	192.5	11.2	744	2 E86255	hypothetical prote
35	192.5	11.2	873	2 A47283	calphotin fruit
36	191.5	11.1	280	2 T11671	extensin-like prot
37	191	11.1	576	2 T36729	probable serine/th
38	190	11.0	550	2 G70597	probable proteinas
39	189.5	11.0	267	2 S08314	cell wall glycopro
40	189	11.0	240	2 B24264	proline-rich prote
41	189	11.0	1794	2 T38459	hypothetical diver
42	188	10.9	544	2 T17547	proline-rich prote
43	188	10.9	699	2 C43674	US4 protein - huma
44	188	10.9	1288	2 T31420	C-terminal domain-
45	187	10.8	350	2 S22456	hydroxyproline-ric

ALIGNMENTS

RESULT 1

D70666
probable modD protein - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70666
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devaiah, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rubner, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70666
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-325 <COL>
A:Cross-references: GB:283859; GB:AL123456; NID:g3261678; PIDN:CAB06127.1; PID:e290722; A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: modD

Query Match 100.0%; Score 1726; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.8e-87;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MHQVDNLTTRKGRLLAALAIAMASNLVTVAVPATANADPEPAPVPTTAASPFSTAAA	60
Db	1	MHQVDNLTTRKGRLLAALAIAMASNLVTVAVPATANADPEPAPVPTTAASPFSTAAA	60
Qy	61	PPAPATPVAPPFAAANTPNAQPGDPNAPPADPNAPPVPIAPNAPQVINDNPVGGF	120
Db	61	PPAPATPVAPPFAAANTPNAQPGDPNAPPADPNAPPVPIAPNAPQVINDNPVGGF	120
Qy	121	SFALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPGPPVANDTRIVLGLDQKLYASAEA	180
Db	121	SFALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPGPPVANDTRIVLGLDQKLYASAEA	180
Qy	181	TDSKAAARLGSDMGFFYMPYPTGTRINQETVSLDANGVSGSASYEYVKESDPSKNGQIWT	240
Db	181	TDSKAAARLGSDMGFFYMPYPTGTRINQETVSLDANGVSGSASYEYVKESDPSKNGQIWT	240
Qy	241	GVIGSPAANAPDAGPPQRFVWLGTANNPVDKGAAKALAESIRPLVAPPAPAPAPAE	300
Db	241	GVIGSPAANAPDAGPPQRFVWLGTANNPVDKGAAKALAESIRPLVAPPAPAPAPAE	300
Qy	301	APAPAGEVAPTPTPTPQRTLPA	325
Db	301	APAPAGEVAPTPTPTPQRTLPA	325

RESULT 2

B87166
 hypothetical protein modd [imported] - Mycobacterium leprae
 C:Species: Mycobacterium leprae
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: B87166
 R: Cole, S.T.; Biglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holt,
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
 A:Title: Massive gene decay in the leprosy bacillus.
 A:Reference number: A86909; PMID:21128732; PMID:11234002
 A:Accession: B87166
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-287 <STO>
 A:Cross-references: GB:AL450380; NID:gl3093663; PIDN:CAC31010.1; GSPDB:GN00147
 C:Genetics:
 A:Gene: modd

Query Match 56.8%; Score 980; DB 2; Length 287;
 Best Local Similarity 66.8%; Pred. No. 8.8e-47;
 Matches 199; Conservative 21; Mismatches 64; Indels 14; Gaps 3;

QY 1 MHQVDNTRKRLAALAAASASLVTVAVPATANADPEAPVPTTAASPPSTAAA 60
 DB 1 MNQVDLSTHRKGLWAILAIAVVASASFTMPLFAAANADPAFL-----PPSTATA 51

QY 61 PPAPATVAPPPAAANTFNAQPGDPAAPPADNAPPVPIAPNAPQPVRIQNPVGGF 120
 DB 52 ASPAQEIITPLGAPVSEHQGPDA--PSLDNAPYPLAVDPNAG--RITNAVGGF 106

QY 121 SPALPAGVSDAAHFDYGSALLSKTTGDPFGPPQPPVANDTRIVLGRDQKLYASAEA 180
 DB 107 SFVLPAWVSEASHLDYGSVLLSKAEQPPVLGQPTVAVTDRIVLGRDQKLYASAEA 166

QY 181 TSKAAALRGSDMGFFYPGTRINQETVSLDANGVSGSASYVEKFSKDPKNGQIWT 240
 DB 167 DNKAALRGSDMGFFYLPYGPTRINQETIPLHANGIAGSASYVEKFSKDPKNGQIWT 226

QY 241 GVIGSPANAPDAGPPQRFVWLGTANNPVDKGAALAESIRPLVAPPPAPAPAPA 298
 DB 227 SVVGSFASSTPDVGSQRFVWLGTNNPVDKGAALAESIRSEMAPIPASVSAPA 284

RESULT 3
 T31611
 hypothetical protein Y50B8A.g - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T31611
 R:Steward, C.
 submitted to the EMBL Data Library, September 1999
 A:Reference number: Z21047
 A:Accession: T31611
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1385 <WIL>
 A:Cross-references: EMBL:AL117200; NID:el549770; PIDN:CAB55050.1; CESP:Y50B8A.g
 A:Experimental source: clone Y50B8A
 C:Genetics:
 A:Gene: CESP:Y50B8A.g
 A:introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match 14.0%; Score 241; DB 2; Length 1585;
 Best Local Similarity 26.6%; Pred. No. 7.1e-06;
 Matches 97; Conservative 12; Mismatches 122; Indels 134; Gaps 13;

QY 41 PEPAPPVPTTAASPPS-----TAAAPPAPATVAPPPAPAA 75
 DB 1233 PPPPPAPAPAPAPSSGGYSGGSGGSAAGGGGGGGYTGSSAAPPFPFPPPPPPAP 1292

QY 76 ANTPNAQPG-----DPNAAPPADPNAPPPPVIAPNAPQVR 112

DB 1293 APAPAFSSGGYSGGSGGSAAGGGGGSGGYSGGSAAPPFPFPPPPPPAPAP-APAP-S 1350
 QY 113 IDNPVGGSFALLPAGWVESDAAHFDYGSALLSKTTGDPFGPPQPPVANDTRIVLGRDQ 172
 DB 1351 SGYSGGSGGSAAGGGGGSGGYTGSSA---APPFPFPPPPPPPPAP----- 1395
 QY 173 KLYASAEATDSKAAARLGSMDGMEFFYPGTRINQETVSLDANGVSGSASYVEKFSKDP 232
 DB 1396 ---APAPAPSSGGYSG-GSSGGS-----AAGGGGGSGGYSGGSAAPP 1434
 QY 233 KNGQIWTGIGSPANAPDAGPPQRFVWLGTANNPVDKGAAXA---LAESIRPLVA 288
 DB 1435 PP-----PPAPAPAPAPSSGGYSG--GSSGGSAAAGGGGGSGGYTGSSAAPP 1478
 QY 289 PPPAP 320
 DB 1479 PPPPPPPPPPPPPAPAPAPAPAPAPAPAPAPSSGGYSGGSGGSAAGGGGGSGGYTGSSAAPPFPFPPPP 1538
 QY 321 RTLPA 325
 DB 1539 PPPPA 1543

RESULT 4
 S49915
 extensin-like protein - maize
 C:Species: Zea mays (maize)
 C:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 29-Oct-1999
 C:Accession: S49915
 R:Rubinstein, A.L.; Broadwater, A.H.; Lowrey, K.; Bedinger, P.A.
 submitted to the EMBL Data Library, June 1994
 A:Description: Pex genes: pollen-specific genes with extensin-like domains.
 A:Reference number: S49915
 A:Accession: S49915
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1188 <RUB>
 A:Cross-references: EMBL:Z34465; NID:g600117; PIDN:CAA84230.1; PID:g600118

Query Match 13.8%; Score 237.5; DB 2; Length 1188;
 Best Local Similarity 28.9%; Pred. No. 8.2e-06;
 Matches 87; Conservative 30; Mismatches 123; Indels 61; Gaps 13;

QY 34 PATANADPEP--APPVPTTAASPPSTAAAPPAPATVAPPPAPAAANTFNAQPGDPAAPP 91
 DB 569 PAPVASPPFPVKSPFPPTLVASPPFPVKSPPPA-PVASPPFPVKSP--PPTPVASPP 624

QY 92 PADPNA-PPPPVIAPNAPQVRIDNPVGGSFALLPAGWVESDAAHFDYGSALLSKTTGDP 150
 DB 625 PPAPVASPPFPVKSP--PPPTPVSSPPFPVKSPPPPPPPPPPPPPPPPPPPPPPPPP 682

QY 151 PFPQPPPPVANDTRIVLGRDQKLYASAEATDSKAAARLGSMDGMEFFYPGTRINQETV 210
 DB 683 PEKSLPPP-----TLISPPPPQEK--FTPTSTPSKPS-----SPEKP 718

QY 211 SLDANGVSGSASYVEKFSKDPKNGQIWTGIGSPANAPDAGPPQRFVWLGTANNP 270
 DB 719 SPKPEPVSSPPQ--TPKSPPPAPAPVSSPPPTPVSSPPALAPVSSPP-----SVKSSP 768

QY 271 VDKGAALAESIRPLVAPPPAP-----APAPAB---PAPAPAPAGVAPPTTPTPTQL 323
 DB 769 PP-----APLSSPPAPAPQVKSPPFPVQVSSFPFPAPKSSPLAPVSSPPQVEKTS 817

QY 324 P 324
 DB 818 P 818

RESULT 5
 S22697
 extensin - Volvox carteri (fragment)
 C:Species: Volvox carteri

Db 28 AAGPVTSTTTAPPTTAAAPTAAAPTPTTTTFVSAQAQPPASPVTPPPAVTPTSPAPKV 87
 QY 76 AN--TPNAQGDNAAPPPADPNAPPPVVIAPNAPQVRIDNPGGFSFALPAGWVESDA 133
 Db 88 AVISAPATPPQPPQPPAGAPTVPSPFPVPPAP----- 122
 QY 134 AHFDYGSALLSKTGPPFPFGQPPPVANDTRIVLGRLDQKLYASAEATDKAAARLGSDM 193
 Db 123 -----TSPPPTPASPPP-----APASPPAPA----- 144
 QY 194 GEFVMPYPTGRINQETVSLDANGVSGSASYEVKFSDPSPKPNQGIWTCVLGSPAANAPDA 253
 Db 145 -----SPPAP-----VSPFPVQAPSP 161
 QY 254 GPPQRFVVMWLTANNFVDKGAALAESIRPLVAPPAPAPAE-----P 300
 Db 162 -----ISLPAPAPAPTCHKRKHKKRHHHAP 198
 QY 301 APAPAPAGEVAPPTPTTPTORTLPA 325
 Db 189 APAPIPPSPSP-FVLTDPQDTAPA 212
 RESULT 7
 T42567
 tegument protein 24 - equine herpesvirus 4 (strain NS80567)
 C:Species: equine herpesvirus 4
 A:Variety: strain NS80567
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
 R:Accession: T42567
 R:Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
 J. Gen. Virol. 79, 1197-1203, 1998
 A:Title: The DNA sequence of equine herpesvirus-4.
 A:Reference number: Z22173; MUID:98264497; PMID:9603335
 A:Accession: T42567
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-3534 <TEL>
 A:Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AACS9539.1; PID:g2605967
 A:Experimental source: strain NS80567
 C:Genetics:
 A:Note: 24
 C:Superfamily: varicella-zoster virus gene 22 protein
 Query Match 13.0%; Score 224.5; DB 2; Length 3534;
 Best Local Similarity 27.8%; Pred No. 0.00012;
 Matches 91; Conservative 18; Mismatches 93; Indels 125; Gaps 13;
 QY 5 DNLTERRKGRLAALAIAMASASLVTVAVPATANADPEPAPVPVTTAASP--PSTAAAPP 62
 Db 2710 DPN----EALLTAPSKPAAPAPSPKAAAPAPSKPAAPAPSKPAAPAPSKPAAPAPSP 2765
 QY 63 APATPVARPPPPAAANTENAPQGDPNAAAPPPADPNAPPPVVIAPNAPQVRIDNPGGFSF 122
 Db 2766 KPAAAPAPSKPAAAPAPS-----KPAAAPAPSKPAAAPAPSKPAAAPAPSK----- 2811
 QY 123 ALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPVANDTRIVLGRLDQKLYASAEATD 182
 Db 2812 --PAA-----APAPSKP-----AAAPAPS 2828
 QY 183 SKAAARLGSDMGCEFFYMPYPTGRINQETVSLDANGVSGSASYEVKFSDPSPKPNQGIWTCV 242
 Db 2829 KPAAA-----PAPSKD----- 2839
 QY 243 IGSPAANAPDAGP-PQRFVVMWLTANNFVDKGAALAESIRPLVAP---PPAPAPAPA 298
 Db 2840 AAAPAPSKPAAAPAPSK-----PAAAPAPSKPAAAPAPS-KPAAAPAPSKPAAAPAPS 2891
 QY 299 EPAPAPAPAGEVAPPTPTTPTORTLPA 325
 Db 2892 KPAAAPAPS-KPAAAPAPSKPNTLVA 2917

Db 91 APATKPTPTAKPATPAPEPPKPPPTPEPKPETPEPPKPAEPKPPKDPPTAEPLKP 150
QY 71 -----PPRAANTPNAGQGNAPADPNAPP-----PPVIANV-----PQ 109
Db 151 VDDTPPPVTPKVTPEVTPKPAFTPEVLOQFVNAQTPPVAKPPVPAPTSTPTPPVQPA 210
QY 110 PVRID-----NPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDP-----PPPGQPPP 158
Db 211 PRTTPPQAARTPNAPAQTPAPATQAPAAQTPPAQAATQTPATPAPAAORPAGAPSP 270
QY 159 VANDTRIVLRDQKLYASAEATDSKAAARLGSMDGEFYMYPGTRI-----NQE 208
Db 271 -----APAPA---QANAPAGSVVPEATVPESSTPAAPSQAQTPPTPRE 310
QY 209 TVSLDANGVSGSASYEYKFSDEPSKPNQIWTGVIGSPAANAPAGPQRFVFWLGTAN 268
Db 311 TAQTEAS-----PAAPNS-----SAAAPNEPASEP-----VAG 338
QY 269 NPVDKGAAKALAEIRPLVAPP-----PAPAPAPAPAPA-----PA 305
Db 339 RP---GTAASGPESASPVTVTPRGETPDTAASAGTPSAGRVTPAPAPSASEGASAARTPG 395
QY 306 PAGEVAPPTPTTP 319
Db 396 AGSQTPPIPATPIP 409

RESULT 11

T29018

hypothetical protein ZK84.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T29018

R:Kirsten, J.

submitted to the EMBL Data Library, April 1995

A:Description: The sequence of C. elegans cosmid ZK84.

A:Reference number: 220553

A:Accession: T29018

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-801 <KIR>

A:Cross-references: EMBL:U23181; PIDN:AAC48204.1; GSPDB:GN00020; CBSP:ZK84.1

A:Experimental source: strain Bristol N2; clone ZK84

C:Genetics:

A:Gene: CBSP:ZK84.1

A:Map position: 2

A:Introns: 22/2; 45/3; 108/1

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Query Match 12.6%; Score 217; DB 2; Length 801;
Best Local Similarity 24.9%; Pred. No. 7.2e-05;
Matches 101; Conservative 37; Mismatches 151; Indels 116; Gaps 16;

QY 25 SASLVTVAPATANADPEAP-----PVPTTAAS---PPST----- 57
Db 194 ASGVVTSEQEGAGDAEPAPVVEETPAPTAAEETAPATEASNAVGTPEGYVDGTANA 253
QY 58 --AAAPAPATPVAPPFAAANTP-NAQPGDNAPPPADPNAPPVPPVIAPNAPQPVRI 114
Db 254 ASAAVEAPVETPA--PAPAAEETAPATSESEAPAPATAETPAPETVSA-APEAANSY 310
QY 115 NPVGGFSFALPAGWVES-----DAAHFDYGSALLSKTTGD----- 149
Db 311 DSAGGDAATAPASSEADAAPDSDASADTTAALVDTSSSEHAESTEAPATDIANTE 370
QY 150 -PPFGGPPPVAN-----DTRIVLGRDQKLYASAEATDSKAAARLGSMDGEF 196
Db 371 TTPAPSVAPVADAAGYDSFSTPEETPAPAAEDTPAPASAAAEETPAPAAAE----- 426
QY 197 YMPVPGTR-----INQETVSLDANGVSGSASYEYKFSDEPSKPNQI 238
Db 427 -TPAPETASAPDAGGAAPVAAADVATTAPETSSAQSAAGSYDV-----PSEAPSE 481

RESULT 13

T06291

extensin homolog T9E8.80 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-Jun-1999

C:Accession: T06291

R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schuell submitted to the Protein Sequence Database, April 1999

A:Reference number: Z15588

QY 239 WTGVIGS-----PAANADAGBPQ--RMFVVVLGTANNPVDK--GAAKALAE 281
Db 482 TAPIVESATPEAPSDSAAPIGFAASEPAPATIEPATDAATLETAPAPAAEPAPAAEAG 541
QY 282 SIRPLVAPPAPAPAP--AEPAPAPAPAGVAPPTTPTTPQRTLPA 325
Db 542 YDAPSSVPETTPAPAPAADETAPAPAAEETPAPAPAAEETPAPA 586

RESULT 12

F96531

hypothetical protein F13F21.7 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: F96531

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: F96531

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-847 <STO>

A:Cross-references: GB:AE005173; NID:G5430752; PIDN:AAD43152.1; GSPDB:GN00141

C:Genetics:

A:Gene: F13F21.7

A:Map position: 1

Query Match 12.5%; Score 216.5; DB 2; Length 847;
Best Local Similarity 26.0%; Pred. No. 8.1e-05;
Matches 79; Conservative 35; Mismatches 135; Indels 55; Gaps 12;

QY 33 VPATANADPEAPPPVPTTAASPPSTAAAPAPATPVAP-----PPAAANTPNAGQCDPNA 88
Db 533 VFPQPPMPSPSPPIYSPPPVHSPPPPVYSSPPPHVYSPPPPVASPPPPPPVPH 592
QY 89 APPPADPNAPPPVIAPNAPQFVRIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTG 148
Db 593 SPPEPPVFSPPPEVTSPPPSPVYSPPPP-----SHSPP-----PPVVS 631
QY 149 DPFPQCPQPPVANDTRIVLG--RLQKLYASAEATDSKAAARLGSMDGEFYMYPVG-TRI 205
Db 632 PPPTTSPPTHTNQPPMGAPTPTCAPTPSETTQVTPSS-ESDQSLSPVQAPTPV 690
QY 206 NOETVSLDANGV---SGSASYEYKFSDEPSKPNQIWTGVIGSPAANAPADAGPPQRFVFW 262
Db 691 QSSTPSSEPTQVPTPSSSEYQAPNLSPVQAPT-----PVQAPTTSETSOVP----- 738
QY 263 WLGTTANNPVDKGAAKALAEIRPLVAPP-----APAPAPAPAPAPAPAGEV-APTPT 315
Db 739 ---TPSSESNQSPQAAPTPILEPVHAPTNSKPVQSPTPS-SEPVSSPEQSSEVEAPEPT 794
QY 316 TPTP 319
Db 795 PVNP 798

RESULT 13

T06291

extensin homolog T9E8.80 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-Jun-1999

C:Accession: T06291

R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schuell submitted to the Protein Sequence Database, April 1999

A:Reference number: Z15588

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 18:17:39 ; Search time 10.5 Seconds
(without alignments)
1611.695 Million cell updates/sec

Title: US-10-720-192-3
Perfect score: 1726
Sequence: 1 MHQVDPNLTGRKRLAALAI.....PAGEVAPTTPTPTQRTILPA 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	1726	100.0	325	1 APA_MYCTU	Q50906 mycobacteri
2	1720	99.7	325	1 APA_MYCBO	Q30620 mycobacteri
3	985	57.1	381	1 APA_MYCAV	Q48919 mycobacteri
4	980	56.8	287	1 APA_MYCLE	P46842 mycobacteri
5	243	14.1	555	1 G11_CHLRE	Q9fpg6 chlamydomon
6	217.5	12.6	817	1 VRP1_YEAST	P37370 saccharomyc
7	213.5	12.4	3164	1 TEGU_HSV11	P10220 herpes simp
8	210	12.2	1083	1 T2D3_HUMAN	Q00268 homo sapien
9	207	12.0	439	1 XP2_XENLA	P17437 xenopus lae
10	205.5	11.9	802	1 ENAH_MOUSE	Q03173 mus musculu
11	203	11.8	3421	1 TEGU_HSV2B	P28955 equine herp
12	197.5	11.4	865	1 CPN_DROME	Q02910 drosophila
13	197	11.4	283	1 EXTN_SORBI	P24152 sorghum bic
14	193	11.2	620	1 EXTN_TOBAC	P13983 nicotiana t
15	192.5	11.2	1198	1 HCN4_RAT	Q9jka7 rattus norv
16	191	11.1	1386	1 ZAP3_MOUSE	Q9r017 mus musculu
17	189.5	11.0	267	1 EXTN_MAIZE	P14918 zea mays (m
18	189	11.0	1794	1 YDC9_SCHPO	Q10172 schizosacch
19	189	11.0	2167	1 SHK1_RAT	Q9wv48 rattus norv
20	188.5	10.9	2715	1 MLL4_HUMAN	Q9umh6 homo sapien
21	188	10.9	699	1 VGLG_HSV2H	P13290 herpes simp
22	186.5	10.8	5179	1 MUC2_HUMAN	Q02817 homo sapien
23	186	10.8	331	1 PRP1_HUMAN	P04280 homo sapien
24	186	10.8	1033	1 IF2_STRCO	Q8cjq8 streptomyce
25	185.5	10.7	1300	1 SAL3_HUMAN	Q9bxa3 homo sapien
26	184.5	10.7	639	1 ZIC5_HUMAN	Q96t25 homo sapien
27	184.5	10.7	1046	1 IF2_STRAW	Q82k53 streptomyce
28	184	10.7	1271	1 RBMG_HUMAN	Q9upn6 homo sapien
29	183	10.6	347	1 CSP_PLABA	P23033 plasmodium
30	182	10.5	261	1 PRP2_MOUSE	P05142 mus musculu
31	180.5	10.5	534	1 APG_ARATH	P40602 arabidopsis
32	180.5	10.5	2716	1 OSA_DROME	Q8in94 drosophila
33	179.5	10.4	2142	1 BAT2_HUMAN	P48634 homo sapien

ALIGNMENTS

RESULT 1

ID	APA_MYCTU	STANDARD	PRT	325 AA
AC	Q50906; Q08062;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Alanine and proline-rich secreted protein apa precursor (45/47 kDa antigen) (Fibronectin attachment protein) (immunogenic protein MPT32)			
DE	(Antigen MPT-32) (45-kDa glycoprotein) (FAP-B).			
GN	APA OR MODD OR RV1860 OR MT1908 OR MTCY359.13.			
OS	Mycobacterium tuberculosis.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
OX	NCBI_TaxID=1773;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=H37RV;			
RC	MEDLINE=96009758; PubMed=7558311;			
RA	Laqueyrie A., Maltizer P., Romain F., Biglmeier K., Cole S.,			
RA	Marchel G.;			
RA	"Cloning, sequencing, and expression of the apa gene coding for the			
RT	Mycobacterium tuberculosis 45/47-kilodalton secreted antigen			
RT	complex."			
RL	Infect. Immun. 63:4003-4010(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=H37RV;			
RC	Laqueyrie A.;			
RA	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=H37RV;			
RC	MEDLINE=98295987; PubMed=9634230;			
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,			
RA	Gordon S.V., Eiglmeier K., Gas S., Barry C.E., III, Tekala F.,			
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,			
RA	Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,			
RA	Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,			
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,			
RA	Rutter S., Seeger K., Skelton S., Squares R.,			
RA	Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;			
RT	"Deciphering the biology of Mycobacterium tuberculosis from the			
RT	complete genome sequence."			
RL	Nature 393:537-544(1998).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CDC 1551 / Oshkosh;			
RC	MEDLINE=22206494; PubMed=12218036;			
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,			
RA	Peterson J.F., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,			
RA	Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,			
RA	Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,			
RA	Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;			
RT	"Whole-genome comparison of Mycobacterium tuberculosis clinical and			
RT	laboratory strains."			

open

J. Bacteriol. 184:5479-5490(2002).
 [5] PARTIAL SEQUENCE, AND CHARACTERIZATION.
 RX MEDLINE=9109989; PubMed=1898899;
 RA Nagai S., Wiker H.G., Harboe M., Kinomoto M.;
 RT "Isolation and partial characterization of major protein antigens in
 RL the culture fluid of Mycobacterium tuberculosis";
 RL Infect. Immun. 59:372-382(1991).
 [6] PARTIAL SEQUENCE, AND GLYCOSYLATION.
 RX MEDLINE=95347792; PubMed=7622204;
 RA Dobos K.M., Swiderik K., Khoo K.-H., Brennan P.J., Belisle J.T.;
 RT "Evidence for glycosylation sites on the 45-kilodalton glycoprotein of
 RL Mycobacterium tuberculosis";
 RL Infect. Immun. 63:2846-2853(1995).
 [7] PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=96196153; PubMed=8626314;
 RA Dobos K.M., Khoo K.-H., Swiderik K.M., Brennan P.J., Belisle J.T.;
 RT "Definition of the full extent of glycosylation of the 45-kilodalton
 RL glycoprotein of Mycobacterium tuberculosis";
 RL J. Bacteriol. 178:2498-2506(1996).
 [8] SEQUENCE OF 40-49, AND CHARACTERIZATION OF CARBOHYDRATE-LINKAGE SITES
 BY MASS SPECTROMETRY.
 RX MEDLINE=H37RV;
 RX MEDLINE=20011399; PubMed=10542234;
 RA Horn C., Namane A., Pescher P., Riviere M., Romain F., Puzo G.,
 Barzu O., Marchal G.;
 RT "Decreased capacity of recombinant 45/47-kDa molecules (Apa) of
 RL Mycobacterium tuberculosis to stimulate T lymphocyte responses
 RT related to changes in their mannosylation pattern";
 RL J. Biol. Chem. 274:32023-32030(1999).
 [9] SEQUENCE OF 40-57, AND CHARACTERIZATION OF CARBOHYDRATE-LINKAGE SITES
 BY MASS SPECTROMETRY.
 RX MEDLINE=H37RV;
 RX MEDLINE=20002527; PubMed=10531201;
 RA Romain F., Horn C., Pescher P., Namane A., Riviere M., Puzo G.,
 Barzu O., Marchal G.;
 RT "Deglycosylation of the 45/47-kilodalton antigen complex of
 RL Mycobacterium tuberculosis decreases its capacity to elicit in vivo
 RT or in vitro cellular immune responses";
 RL Infect. Immun. 67:5567-5572(1999).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MASS SPECTROMETRY: MW=28782; MW_ERR=1.6; METHOD=Electrospray;
 CC RANGE=40-325.
 CC -1- BIOTECHNOLOGY: Major immunodominant antigen that has potential as
 CC a vaccine against tuberculosis. APA-ELISA could be used in
 CC diagnosis.
 CC -1- MISCELLANEOUS: Changes in the mannosylation pattern of this
 CC protein affect its ability to stimulate T-lymphocyte response.
 CC -1- CAUTION: Was originally thought to be involved in molybdenum
 CC transport.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X80268; CAA56555.1; --
 DR EMBL; X99258; CAA67645.1; --
 DR EMBL; Z83859; CAB06127.1; --
 DR EMBL; AE007048; AAK46179.1; --
 DR PIR; D70666; D70666.
 DR TIGR; MT1908; --
 DR TubercuList; Rv1860; --
 KW Antigen; Glycoprotein; Repeat; Signal; Complete proteome.
 FT SIGNAL 1 329
 FT CHAIN 40 325 ALANINE AND PROLINE-RICH SECRETED PROTEIN

FT DOMAIN 85 107 APA.
 FT REPEAT 85 88 N-A.
 FT REPEAT 94 97 1.
 FT REPEAT 104 107 2.
 FT CARBOHYD 49 49 O-LINKED (MAN. . .).
 FT CARBOHYD 57 57 O-LINKED (MAN. . .).
 FT CARBOHYD 66 66 O-LINKED (MAN. . .).
 FT CARBOHYD 316 316 O-LINKED (MAN. . .).
 FT CONFLICT 136 136 F -> L (IN REF. 4).
 SQ SEQUENCE 325 AA; 32720 MW; 59E5D0455A997BED CRC64;
 Query Match 100.0%; Score 1726; DB 1; Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.5e-80;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MHQVDNLTERRKRLAALAIAAASASLVTVAVPATANADPEAPVPTTAASPPSTAAA 60
 DB 1 MHQVDNLTERRKRLAALAIAAASASLVTVAVPATANADPEAPVPTTAASPPSTAAA 60
 QY 61 PPAPATVPAPPPAAANTNAQPGDPAADNPAPPPVIAAPNAPQVRIAPVGGF 120
 DB 61 PPAPATVPAPPPAAANTNAQPGDPAADNPAPPPVIAAPNAPQVRIAPVGGF 120
 QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFPGPPVANDTRIVLGRDQKLYASAEA 180
 DB 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFPGPPVANDTRIVLGRDQKLYASAEA 180
 QY 181 TDSKAAARLGSDMGFYMYPGTRINQETVSLDANGVSGSASYEYKFSDPKNGQIWT 240
 DB 181 TDSKAAARLGSDMGFYMYPGTRINQETVSLDANGVSGSASYEYKFSDPKNGQIWT 240
 QY 241 GVIGSPAANAPDAGPPQRFVWLGTANNPVDKGAALAEISIRPLVAPPPAPAPAPAP 300
 DB 241 GVIGSPAANAPDAGPPQRFVWLGTANNPVDKGAALAEISIRPLVAPPPAPAPAPAP 300
 QY 301 APAPAPAGEVAPTPTTPTPTQRTLPA 325
 DB 301 APAPAPAGEVAPTPTTPTPTQRTLPA 325
 RESULT 2
 APA_MYCBO STANDARD; PRT; 325 AA.
 AC O30620;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Alanine and proline-rich secreted protein apa precursor (Fibronectin
 DE attachment protein) (45/47 kDa antigen) (FAP-B).
 GN APA OR MB1891.
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_TaxID=1765;
 RX SEQUENCE FROM N.A.
 RN STRAIN=BCG;
 RA Zhao W., Schorey J.S., Bong-Mastek M., Brown E.J., Ratliff T.L.;
 RT "Identification, sequence and characterization of the M. bovis BCG
 RT fibronectin attachment protein";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972;
 RA Gardner T., Eigler K., Camus J.-C., Medina N., Mansoor H.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrall B.G., Cole S.F., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).


```
Db 263 APPPPPPPPPPPP-----RP-----PPFA-----NTPMPSPSPSP 293
QY 157 PPVANDTRIVLGRDLQKLYASAEATDSKAAARLGSMDGMEFYMPGTRINQETVSLDANG 216
Db 294 PB-----SPAPPT-----301
QY 217 VSGSASYEVKFSKNGKQIWTGVIQSPAAAPDAGPQRWFWVWLGTANNPVDKGA 276
Db 302 -----PTTSPSPSPSPVPVPPSPAPVPPSPAPPS-----329
QY 277 KALAESIRPLVAPPAPAPAPAPAPAPAGEVAPTPT-TPTP 319
Db 330 -----PAPSPSPSPAPPTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 365
RESULT 6
VRP1 YEAST
ID VRP1 YEAST STANDARD; PRT; 817 AA.
AC P37370; Q06133;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Verprolin.
GN VRP1 OR MP2 OR END5 OR YLR337W OR L8300.13.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A364;
RX MEDLINE=95058201; PubMed=7968536;
RA Donnelly S.F.H., Pocklington M.J., Pallota D., Orr E.;
RT "A proline-rich protein, verprolin, involved in cytoskeletal
RT organization and cellular growth in the yeast Saccharomycetes
RT cerevisiae.";
RL Mol. Microbiol. 10:585-596(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston W., Hillier L., Riles L., Albertmann K., Andre B., Anserge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoft A.,
RA Enkian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koelter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Netwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetalle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997)
CC -!- FUNCTION: Involved in cytoskeletal organization and cellular
CC growth. May exert its effects on the cytoskeleton directly, or
CC indirectly via proline-binding proteins (e.g. profilin) or
CC proteins possessing SH3 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; Z26645; CAB81388.1; -
CC DR EMBL; U19028; AAB67263.1; -
CC DR PIR; S51342; S51342.
CC DR GenOnline; 142401; -
CC DR SGD; S0004329; VRP1.
CC GO; GO:0005857; C:actin cortical patch (sensu Saccharomycetes); IDA.
```

```
DR GO; GO:0003779; F:actin binding; IDA.
DR GO; GO:0007015; P:actin filament organization; IPI.
DR GO; GO:0006997; P:endocytosis; IMP.
DR GO; GO:0007121; P:polar budding; IMP.
DR GO; GO:0006370; P:response to osmotic stress; IMP.
DR InterPro; IPR003124; WH2.
DR Pfam; PF02295; WH2; 2.
DR SMART; SM00246; WH2; 2.
KW Cytoskeleton; Repeat.
FT DOMAIN 5 14 POLY-PRO.
FT DOMAIN 239 245 POLY-PRO.
FT DOMAIN 349 357 POLY-PRO.
FT DOMAIN 396 406 POLY-PRO.
FT DOMAIN 424 431 POLY-PRO.
FT DOMAIN 462 468 POLY-SER.
FT DOMAIN 704 708 POLY-PRO.
FT CONFLICT 308 308 P -> R (IN REF. 1).
FT CONFLICT 350 350 A -> R (IN REF. 1).
FT CONFLICT 689 689 V -> E (IN REF. 1).
FT CONFLICT 710 817 PSTMDTGTSTNSPSKNLQKRLFFSTGGSTLQHKHNTHTNQDPV
FT DVGRYTGSSNIVGAKSGNERIVDDISGRKWTNVNQMPKP
FT RPFQNTKLYPSGKSGSVFLDLFT -> HLEWIPVPLIA
FT PVKTLNNGVFLQVDRRCNTSIIRIQINQMLM (IN REF.
FT 1).
SQ SEQUENCE 817 AA; 82593 MW; 24C752D5B1CA1C8 CRC64;
Query Match 12.6%; Score 217.5; DB 1; Length 817;
Best Local Similarity 26.3%; Pred. No. 0.00018;
Matches 99; Conservative 31; Mismatches 125; Indels 121; Gaps 18;
QY 24 ASASLVTVAVPATA-----NADPEAPPVPT--TAASPP-----STAAAPPAPATPVAPP 72
Db 118 ASAPPITGAVPSVAAPPINAPLSPAPAVPSIPSSSAPPIDIPSSAAPPPIVPSPPAP 177
QY 73 P-----AAANTNAQFGDPNAAP-----PPA-----DPN 96
Db 178 FLPLSGASAPKVPQNRPHMPVRPAHRSHQKSSNISLPSVSAPPLPSASLPTHVSNPPQ 237
QY 97 APPPPV-----IAP--NAPQVRIDNPVGGFSEALPAGWSDSDAHPDYG-----140
Db 238 APPPTPTTIGLDSKNIKPTDNAVSPSSEVPAGGLPFL-----ABINARRSERGAVEGV 292
QY 141 -----ALLSKTTGDPFFGQPPPVAND-----TRIVLGRDLQKLYASAEATDSKAAA 187
Db 293 SSTKIQTENHKSPQPPFPSSAPPIPTSHAPPLPTAPPPLPNVTSAPKKATSAPAP 352
QY 188 RLQSDMGFFMPYPGTRINQETVSLDANGVSGSASYEVKFSKNGKQIWTGVIQSPA 247
Db 353 P-----PPLPAAVSSASTNSVKATVPPTLA-----PPLPN---TTSVPPNKA 392
QY 248 ANAPDAGPPQRWFWVWLGTANNPVDKGA---AKALAESIRPLVAPPAPAPAPAPAP-A 303
Db 393 SSMPPPPPP-----PPPGAPSTSSALSASSIPLAPLPFPPPPSVATSVPSA 440
QY 304 PAPAGEVAPTPTPTPT 319
Db 441 PPP-----PPTLTATKP 452
RESULT 7
TEGU_HSV11
ID TEGU_HSV11 STANDARD; PRT; 3164 AA.
AC P10220;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Large tegument protein (Varion protein UL36).
GN UL36.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
```

SEQUENCE FROM N.A.
MEDLINE=88274327; PubMed=2839594;
McGeoch D.J., Dalrymple M.A., Davidson A.J., Dolan A., Frame M.C.,
McNab D., Perry L.J., Scott J.E., Taylor P.;
"The complete DNA sequence of the long unique region in the genome of
herpes simplex virus type 1";
J. Gen. Virol. 69:1531-1574 (1988).
-!- FUNCTION: Tegument protein.
-!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
EHV-1 24, EBV BFLF1, HVS-1 64, VZV 22, AND HCMV UL48.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X14112; CAA32311.1; -
PIR; I30085; WMBH6.
InterPro; IPR006928; Herpes teg N.
InterPro; IPR005210; Herpes UL36.
Pfam; PF04843; Herpes teg N1.
Pfam; PF03586; Herpes_UL36; 1.
Repeat.
DOMAIN 2911 2980 35 X 2 AA TANDEM REPEATS OF P-Q.
SEQUENCE 3164 AA; 335857 MW; CC5D31FF4F9FE3F4 CRC64;

Query Match 12.4%; Score 213.5; DB 1; Length 3164;
Best Local Similarity 26.5%; Pred. No. 0.001;
Matches 91; Conservative 17; Mismatches 117; Indels 119; Gaps 15;

QY 10 RKGRALALAIAMASA-----SLTVAVPATANADPEAPVPVPTAASPSTAAAP 61
DB 2683 RRRHRRSLARATQASATTQGRPPALPDVTAVPTDFARP-PAPKPPPEPAPHALVSGVP 2741

QY 62 -----PAPATVAP-PPPAANTNAGDPENAPPPADNPAPPPVADNAPQP 110
DB 2742 LPLGQAQAGALPIDVPPVPAVTG--VLPGCENRRPPLTSPATPRPVGGQR 2799

QY 111 VRIDNPVGGFSPALPAGWVEDAAHFDYGSALLSKTGDPPPPGPPVANDTRVLGLR 170
DB 2800 -RLTRP-----AVASLSSESRESLPSFMDP---ADPTAPVLGR- 2832

QY 171 DQKLVASAATSKAARLGDGMEFYMPYPCPTRINQETVSLDANGVSSASYYEVKFS 230
DB 2833 -----NPAPETSSPAG----- 2844

QY 231 PSKPNQIWTGIVGSPAANAPDAGPPQRFVVLGTA-----NNPVDKGAALAESIR 284
DB 2845 PSPPPAV-----QVAPPTSGPPTVLTLEGGVAPGVPVSRRTTRQPVATPTTSAR 2898

QY 285 P-----LVAPPAPAPAPAPAPAPAPAGEVAPTPTTTPQ 320
DB 2899 PRGHLTVSLRSLAPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 2940

RESULT 8
T2D3_HUMAN
ID T2D3_HUMAN STANDARD; PRT; 1063 AA.
AC Q00268; Q99721; Q9BR40; Q9BX42;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Transcription initiation factor TFIID 135 kDa subunit (TAFII-135)
DE (TAFII135) (TAFII130) (TAFII130)
GN TAF4 OR TAF4A OR TAF2C1 OR TAF2C OR TAFII135 OR TAFII130.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

SEQUENCE FROM N.A.
MEDLINE=97336072; PubMed=9192867;
Mengus G., May M., Carre L., Chambon P., Davidson I.;
"Human TAF(II)135 potentiates transcriptional activation by the AF-2s
of the retinoic acid, vitamin D3, and thyroid hormone receptors in
mamalian cells";
Genes Dev. 11:1381-1395 (1997).
[2]
SEQUENCE FROM N.A.
MEDLINE=21638749; PubMed=11780052;
Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,
Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.F., Clegg S.,
Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
Marsh V.L., Martin S.L., McConnachie L.J., McLeay K., McMurray A.A.,
Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
Rogers J.;
"The DNA sequence and comparative analysis of human chromosome 20";
Nature 414:865-871 (2001).
[3]
SEQUENCE OF 105-1083 FROM N.A., AND PARTIAL SEQUENCE.
MEDLINE=97098442; PubMed=8942982;
Tanese N., Saluja D., Vassallo M.F., Chen J.-L., Admon A.;
"Molecular cloning and analysis of two subunits of the human TFIID
complex: hTAFII130 and hTAFII100";
Proc. Natl. Acad. Sci. U.S.A. 93:13611-13616 (1996).
[4]
IDENTIFICATION IN THE TFIID-HAT COMPLEX WITH TAF5L; TAF6L; TADA3L;
SUPT3H; TAF2; TAF5; TTRAP; GCN5L2 AND TAF10.
MEDLINE=99303588; PubMed=10373431;
Brand M., Yamamoto K., Staub A., Tora L.;
"Identification of TAF4-binding protein-free TAFII-containing complex
subunits suggests a role in nucleosome acetylation and signal
transduction";
J. Biol. Chem. 274:18285-18289 (1999).
[5]
X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 870-918 IN COMPLEX WITH
TAF12.
MEDLINE=20063193; PubMed=10594036;
Gangloff Y.-G., Werten S., Romier C., Carre L., Poch O., Moras D.,
Davidson I.;
"The human TFIID components TAF(II)135 and TAF(II)20 and the yeast
SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like
pairs";
Mol. Cell. Biol. 20:340-351 (2000).
-!- FUNCTION: Makes part of TFIID is a multimeric protein complex that
plays a central role in mediating promoter responses to various
activators and repressors. Potentiates transcriptional activation
by the AF-2s of the retinoic acid, vitamin D3 and thyroid hormone.
-!- SUBUNIT: TFIID is composed of TATA binding protein (TBP) and a
number of TBP-associated factors (TAFs). Component of the TFIID-HAT
complex, at least composed of TAF5L, TAF6L, TADA3L, SUPT3H/SPT3,
TAF2/TAFII150, TAF4/TAFII135, TAF5/TAFII100, GCN5L2/GCN5, TAF10
and TTRAP.
-!- SUBCELLULAR LOCATION: Nuclear.

```

CC  -!- SIMILARITY: BELONGS TO THE TAF2C FAMILY.
CC  -!- SIMILARITY: Contains 1 TAFH/NHR1 domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; Y11354; CAA72189.1; -.
DR  EMBL; AL137077; CAC36006.1; -.
DR  EMBL; AL109911; CAC22312.2; -.
DR  EMBL; U75308; AAC50901.1; -.
DR  PDB; 1H3O; 26-SEP-02.
DR  TRASNAC; T02328; -.
DR  Genew; HGNC:11537; TAF4.
DR  MIM; 601796; -.
DR  GO; GO:0005669; C:transcription factor TFIID complex; TAS.
DR  GO; GO:0016251; F:general RNA polymerase II transcription fac. .; TAS.
DR  GO; GO:0005515; F:protein binding; TAS.
DR  GO; GO:0003713; F:transcription co-activator activity; TAS.
DR  InterPro; IPR007900; TAF4.
DR  InterPro; IPR003894; TAF_hom.
DR  Pfam; PF05236; TAF4; 1.
DR  SMART; SM00549; TAFH; 1.
DR  Transcription regulation; Nuclear protein; 3D-structure.
FT  DOMAIN 590 682
FT  TAFH/NHR1.
FT  DOMAIN 39 42
FT  POLY-HIS.
FT  DOMAIN 52 57
FT  POLY-ALA.
FT  DOMAIN 98 101
FT  POLY-GLY.
FT  DOMAIN 142 148
FT  POLY-ALA.
FT  DOMAIN 268 275
FT  POLY-PRO.
FT  DOMAIN 331 337
FT  POLY-ALA.
FT  DOMAIN 680 683
FT  POLY-PRO.
FT  DOMAIN 808 813
FT  POLY-ALA.
FT  DOMAIN 828 831
FT  POLY-ASP.
FT  CONFLICT 105 117
FT  PGSPSPRLVPA -> GRGLLQORGGRES
FT  (IN REF. 3).
FT  CONFLICT 136 136
FT  A -> S (IN REF. 2).
FT  CONFLICT 185 185
FT  G -> GPG (IN REF. 2).
FT  CONFLICT 233 264
FT  MISSING (IN REF. 3).
FT  CONFLICT 293 293
FT  P -> L (IN REF. 3).
SQ  SEQUENCE 1083 AA; 109943 MW; A6453827572A0752 CRC64;

Query Match 12.2%; Score 210; DB 1; Length 1083;
Best Local Similarity 27.8%; Pred. No. 0.00056;
Matches 97; Conservative 20; Mismatches 140; Indels 92; Gaps 13;

Qy 15 LAALAIAMAGASLVTVAVPATANADPEAPVPVTTAASPSTAAAPAPATPVAPPPPA 74
Db 242 VGLAAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 300
Qy 75 AAN--TPNAQCPDNAAAPPADPNAPPVPPVPIAPNAPOF-VRIDNPVGFGFALPAGWV-- 129
Db 301 AONGSAGAPAPAPAGGAGVQPGGPGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Qy 130 ----ESDAAHFDYG-----SALLSKTGTGDPFGPPGPPVANDTRIVGLDQKLYASAEAT 181
Db 361 ASGPASTAASMVIGTQWGLPSPAAPPAPPPAGPTGLPKG---AAGAVTQSLSRTPAT 417
Qy 182 DSKAAARLGDWGFYMPYPGTRINQETVSLDANGVSGSASYEVKFSDFSKP-----NGQ 237
Db 418 TSGIRATLTPTVLAFLPQP-----PQNPNINQNFQ 448
Qy 238 IWTGVISPAANADAGSPQRFVWLGTANNPVVDKGAALAE-STRPLVAPPAPAPA 296
Db 449 LPPGGLVLRSENGOLLMIPO-----QALAQQAQAHAQAPOTTMAPR 489
Qy 297 PAEPAPAP-----AP-----AGEVAPT-----PTTPTPQRTL 323
Db 490 PATTSAPPVQISTVQAPGTPIIARQVTPTTIHKVQSAQTIVQPSATL 538

```

RESULT 9

```

XP2_XENIA STANDARD; PRT; 439 AA.
ID XP2_XENIA
AC P17437; Q08944;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Skin secretory protein XP2 precursor (APSG protein).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE OF 1-25 AND 344-439 FROM N.A. (ISOFORM 2).
RX MEDLINE=9232564; PubMed=1629230;
RA Hauser F., Roeben C., Hoffmann W.;
RT "XP2, a new member of the P-domain peptide family of potential growth
RT factors, is synthesized in Xenopus laevis skin.";
RL J. Biol. Chem. 267:14451-14455(1992).
RN [2]
RP SEQUENCE OF 3-439 FROM N.A. (ISOFORM 1).
RC TISSUE=Skin;
RX MEDLINE=90127399; PubMed=2298293;
RA Gmachl M., Berger H., Thalhammer J., Kreil G.;
RT "Dermal glands of Xenopus laevis contain a polypeptide with a highly
RT repetitive amino acid sequence.";
RL FEBS Lett. 260:145-148(1990).
CC -!- FUNCTION: May act as a growth factor in the germinal layer of the
CC epidermis. May also be involved in growth of regenerating glands
CC and in protection of the skin from the external environment.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=APEG;
CC IsoId=P17437-1; Sequence=Displayed;
CC Name=2; Synonyms=XP2;
CC IsoId=P17437-2; Sequence=VSP_004652;
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: Contains 2 P-type (trefoil) domains.
CC -!- CAUTION: Ref.2 sequence differs from that shown from position 392
CC onward and is shorter (418 AA) due to a frameshift.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; M90095; AAA50001.1; -.
DR  EMBL; X51394; CAA35759.1; ALT_FRAME.
DR  PIR; A37331; A37331.
DR  PIR; S07498; SKXLAG.
DR  HSSP; P04155; 1PS2.
DR  InterPro; IPR000519; P_trefoil.
DR  Pfam; PF00088; trefoil; 2.
DR  PRINTS; PRO0680; PTREFOIL.
DR  SMART; SM00018; PD; 2.
DR  PROSITE; PS00025; P_TREFOIL; 2.
DR  Signal; Growth factor; Alternative splicing; Repeat;
DR  Pyroliidone carboxylic acid.
KW  SIGNAL 1 22 POTENTIAL.
FT  CHAIN 23 439 SKIN SECRETORY PROTEIN XP2.
FT  MOD_RES 23 23 PYRROLIDONE CARBOXYLIC ACID (PROBABLE).
FT  DOMAIN 26 343 33 X REPEATS OF G-[GE]-[AP] (2,4)-A-E.
FT  DOMAIN 350 391 P-TYPE 1.
FT  DOMAIN 397 438 P-TYPE 2.
FT  DISULFID 351 377 BY SIMILARITY.
FT  DISULFID 361 376 BY SIMILARITY.

```



```

FT  SEQUENCE 802 AA; 85844 MW; 592BB3975EE20F7F CRC64;
SQ  QUERY Match 11.9%; Score 205.5; DB 1; Length 802;
    Best Local Similarity 22.9%; Pred. No. 0.00072;
    Matches 78; Conservative 27; Mismatches 112; Indels 123; Gaps 11;

QY 33 VPATANADPEAPPPVTTAAAPPSTAA-----PPAPATPVADPP----- 72
DB 343 VPRUNKRSRSPSYNTSSQPPAKSCAWPTNSFSPPLPSPPIMISSPPGKATGPRVL 402
QY 73 PAAANTPNAQ-----PGDNNAA-----PPADENAPPPVPIAPNAPOVRIDNPVG 118
DB 403 PUCVSSPVQPPSPSTAENGSLDSVTYFVSPPTSGRAAPPPPPPPPPPP----- 454
QY 119 GFSFALPAGVSDAHDYGSALLSKTTGPPPPGQP-PPVANDTRIVLGRLOKLYAS 177
DB 455 -----PPLPPPLPPLAS----- 467
QY 178 AEATDSKAAARLGSQWGEFYMYPCTRI-----NQETVSLDANGVSGSAS---YVEVKF 228
DB 468 -----LSHCGSQAPPPCTPLASTPSSKSPVLPSPSAGAPASATPLNPELGD 515
QY 229 SDPSKPNQIWTGVTGSPAANADAGPPQRFVWLGTANNPVDKGAALAEISIRPLVA 288
DB 516 SSASEP-----GLQAASQPAESPTPQGLVLGPPAPPPPPPLPSGPAVASALPPPPG 566
QY 289 ---PPAPAPAPAEAPAPAPAGEVAPTPTTPTPTPLPA 325
DB 567 PPPPPPLPSTGPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 606

RESULT 11
ID TEGU_HSVB STANDARD; PRT; 3421 AA.
AC F28955;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Large tegument protein.
GN 24.
OS Equine herpesvirus type 1 (strain Ab4p) (HEV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1."
RL Virology 189:304-316(1992).
CC -!- FUNCTION: Tegument protein.
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
    HSV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
    the European Bioinformatics Institute. There are no restrictions on its
    use by non-profit institutions as long as its content is in no way
    modified and this statement is not removed. Usage by and for commercial
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
    or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M86664; AA02459.1; -.
DR PIR; G36797; WZBEE6.
DR InterPro; IPR006928; Herpes teg N.
DR InterPro; IPR005210; Herpes UL36.
DR Pfam; PF04843; Herpes teg N; 1.
DR Pfam; PF03586; Herpes_UL36; 1.
SQ SEQUENCE 3421 AA; 367078 MW; 5075EPE4739BB7AC CRC64;

Query Match 11.8%; Score 203; DB 1; Length 3421;
Best Local Similarity 23.9%; Pred. No. 0.0036;

Matches 96; Conservative 31; Mismatches 152; Indels 122; Gaps 15;

QY 6 PNLTRKGRALALAIAMASASLIVTAVPATANADPEAPPPVTTAAAPP-----STAAA 60
DB 2505 PTLPPKAAAPLPPSDASAIMSGKPVFKYITGKNSAVPSPAPPTLPPAPLPQSTSKAAS 2564
QY 61 PPAPATPVADPPPPAAATNAQPGDNNAAAPPADPNAP-----PPPVIAFNAP 108
DB 2565 GPPTLPPAPPLP-OSTSKAASG-----PPPTLPPAPLPQSTSKAASGPPPTLPPAPP 2617
QY 109 QPVRIIDNPVGGF-----SFALPAGWVESDAA--- 134
DB 2618 LPOSTSKAASGATQSDSGKTLTLDVKTSQKQVFPVPTDKPSTTTPAALQSDASKPP 2677
QY 135 -----HFDYGSALLSKTTGDDPP-----FPG-QPP-----PVANTRIV 166
DB 2678 TAAIQHQKLGITVTPTKSGDKFTDNASAFVGVSPVTPDGTGAKPFPKADPPVDTK-- 2735
QY 167 LGRLDOKLYASAEATDSKAAARLGSQWGEF-YMPYPTTRINQETVSLDANGVSGSASYE 225
DB 2736 -----QPVKSLSPQVRGPPYIRPSLGFKFTGPEG-----YTIPIVHGLPPSDSNVT 2783
QY 226 VKFSDPSKPNQIWTGVTGSPAANADAGPPQRFVWLGTANNPVDKGAALAEISIRP 285
DB 2784 QSTKEPPK-----PAVETPAAPAK-----SAAAPAAAPAKSAAAPAAAP 2823
QY 286 L-VAPAPAPAPAEAPAPAPAGEVAPTPTTPTPTPLPA 325
DB 2824 AKSAAAPAAAPAKSAAAPAAAPAKSAAAPAAAPAKDQTKSA 2864

RESULT 12
ID CPN_DROME STANDARD; PRT; 865 AA.
AC Q02910;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Calphotin.
GN CPN OR CAP.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Canton-S;
RX MEDLINE=93185729; PubMed=8094559;
RA Martin J.H., Benzer S., Rudnicka M., Miller C.A.;
RT "Calphotin: a Drosophila photoreceptor cell calcium-binding protein."
RL Proc. Natl. Acad. Sci. U.S.A. 90:1531-1535(1993).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=Canton-S;
CC MEDLINE=93185730; PubMed=8434015;
CC RA Ballinger D.G., Xue N., Harshman K.D.;
CC RT "A Drosophila photoreceptor cell-specific protein, calphotin, binds
    calcium and contains a leucine zipper."
CC Proc. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).
CC -!- FUNCTION: Might function as a calcium-sequestering "sponge" to
    regulate the amount of free cytoplasmic calcium. It binds 0.3 mole
    of Ca(2+) per mole of protein.
CC -!- SUBUNIT: Homodimer (Probable).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; hypodense compartment.
CC -!- TISSUE SPECIFICITY: Soma and axons of photoreceptor cells of
    compound eyes and ocelli.
CC -!- DEVELOPMENTAL STAGE: Expressed early in photoreceptor cell
    development.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
    the European Bioinformatics Institute. There are no restrictions on its

```


KW Potassium channel; Potassium transport; Sodium transport;
KW CAMP; cAMP-binding; Transmembrane; Glycoprotein; Sodium channel.
FT DOMAIN 1 286 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 267 287 SEGMENT S1 (POTENTIAL).
FT TRANSMEM 294 314 SEGMENT S2 (POTENTIAL).
FT DOMAIN 315 340 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 341 361 SEGMENT S3 (POTENTIAL).
FT TRANSMEM 369 389 SEGMENT S4 (POTENTIAL).
FT DOMAIN 390 420 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 421 441 SEGMENT S5 (POTENTIAL).
FT TRANSMEM 465 486 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
FT TRANSMEM 497 517 SEGMENT S6 (POTENTIAL).
FT DOMAIN 518 1198 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 209 260 INVOLVED IN SUBUNIT ASSEMBLY (BY
FT SIMILARITY).
FT NP BIND 595 710 CAMP.
FT DOMAIN 799 1074 PRO/SER-RICH.
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 404 404 I -> V (IN REF. 2).
SQ SEQUENCE 1198 AA; 128760 MW; 6B92B8F9452F760F CRC64;

Query Match 11.2%; Score 192.5; DB 1; Length 1198;
Best Local Similarity 23.8%; Pred. No. 0.0047;
Matches 98; Conservative 37; Mismatches 124; Indels 153; Gaps 19;

QY 10 RKGRLLAALAIAMASASLV-----VAVPATA-----NADPEPAPVPTTAASPP 55
Db 824 RHPRRLQSLIPALGASGASPASSQVDTPSSSSFIQQLAGFSAPPGLSPLLPSSSSSPP 883
QY 56 STA-AAPPAPA-----TPVAP-----70
Db 884 PGACSSPPAPTPTSTAATTGTFGHFKALGSLSSSDSPLLTLPQCARSPQAAPPPP 943
QY 71 -----PPPAANTPNAQG-----DPN-AAPPPADPNAPPPP-----101
Db 944 LPGARGGLLHFLPPPPSSPSPPGQLQCPGELSPGLAGPPSTPTPTPRPSPS 1003
QY 102 ---VIAPNAPQVRIQNDPVGGFSFALPAGWVESDAHF-----DYGSAALLSKTYGDP 150
Db 1004 FMAGASGGASPVAF-TPRGGLS---PPGHSFGPPRTFPSPAPPASGSHGSLLL-----P 1053
QY 151 PFFGQPPPVANDTR-----IVLGRLDQKL---YASAEATDSKAAEL-----GSDMGE 195
Db 1054 PASSPPPPQVPPORRTPTLTGRLTQDLKLSASQFALPDQCAQTLRRASPHSGESVAA 1113
QY 196 FYMPYFGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWTGVIGSPAANAPDAGP 255
Db 1114 FSL-YF-----RAGGSGSGG---GLGPPGRPYGAI PQQHVTLPRKTSGLSLP 1157
QY 256 PQRFVFWVLGTANNPVDKGAKAKLAESIRPLVAPPAPAPAPAPAPAPAPA 307
Db 1158 P-----PLSLFGAKAASSGGPPLTAAPQREPGARSEFVRSKLPS 1196

Search completed: July 7, 2004, 18:21:52
Job time : 11.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 18:18:15 ; Search time 36.5 seconds
(without alignments)
2809.409 Million cell updates/sec

Title: US-10-720-192-3
Perfect score: 1726
Sequence: 1 MHQVDPNLTTRKGRALALAI.....PAGEVAPTPTTPTPQRTLPA 325

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_invertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriopl.*
17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1539	89.2	286	Q9r640	mycobacteri
2	1053.5	61.0	368	Q9r4H9	mycobacteri
3	568	32.9	194	Q8V565	mycobacteri
4	256	14.8	3247	Q65553	bovine herp
5	248.5	14.4	698	Q9ASK4	oryza sativ
6	248	14.4	3288	Q7T5D9	simian herp
7	244	14.1	511	Q95JD0	sus scrofa
8	244	14.1	566	Q95JD1	sus scrofa
9	241.5	14.0	745	Q89X06	bradyrhizob
10	240.5	13.9	3326	Q7T591	simian herp
11	238	13.8	1269	Q8W5K6	oryza sativ
12	238	13.8	1269	Q7XH56	oryza sativ
13	237.5	13.8	1188	Q41805	zea mays (m
14	231	13.4	668	Q8XU08	caenorhabdi
15	230.5	13.4	584	Q82DU7	streptomyce
16	228.5	13.2	464	Q41645	volvox cart

17	227.5	13.2	552	4	Q8NAA1	Q8naal homo sapien
18	227	13.2	598	16	Q8VKN7	Q8vkn7 mycobacteri
19	226.5	13.1	222	10	Q9S740	Q9s740 arabidopsis
20	226.5	13.1	609	16	Q8FM53	Q8fm53 corynebacte
21	225.5	13.1	420	5	Q9VZC2	Q9vzc2 drosophila
22	225.5	13.1	652	10	Q94JZ6	Q94jz6 arabidopsis
23	225	13.0	676	6	Q95JC9	Q95jc9 sus scrofa
24	224.5	13.0	3534	12	Q39266	Q39266 equine herp
25	223.5	12.9	652	10	Q91V48	Q91v48 arabidopsis
26	222	12.9	763	2	Q9XDH2	Q9xdh2 mycobacteri
27	222	12.9	3084	12	Q8UZ11	Q8uz11 pseudorabie
28	218	12.6	409	10	Q8SBM1	Q8sbm1 volvox cart
29	217.5	12.6	839	16	Q9RX57	Q9rx57 deinococcus
30	217	12.6	784	13	Q90YB5	Q90yb5 gallus gall
31	217	12.6	801	5	Q23635	Q23635 caenorhabdi
32	216.5	12.5	496	16	Q7U8L8	Q7u8l8 synechococc
33	216.5	12.5	847	10	Q9XIB6	Q9xib6 arabidopsis
34	215.5	12.5	817	3	Q07229	Q07229 saccharomyc
35	214.5	12.4	687	10	Q948Y7	Q948y7 volvox cart
36	214.5	12.4	1016	10	Q9SPM0	Q9spm0 zea mays (m
37	214	12.4	1204	3	Q8JOE5	Q8joe5 yarrowia li
38	214	12.4	1627	10	Q84ZL0	Q84zl0 oryza sativ
39	213.5	12.4	760	10	Q9TOK5	Q9tok5 arabidopsis
40	213.5	12.4	3122	12	P89459	P89459 herpes simp
41	212	12.3	753	5	Q93107	Q93107 acanthamoeb
42	211	12.2	889	16	Q9F2N5	Q9f2n5 streptomyce
43	209	12.1	1480	4	Q96Q04	Q96q04 homo sapien
44	208.5	12.1	214	10	Q93763	Q93763 gossypium b
45	206.5	12.0	575	16	Q98F98	Q98f98 rhizobium l

ALIGNMENTS

RESULT 1

ID Q9R640 PRELIMINARY; PRT; 286 AA.
AC Q9R640;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MPT-32=45 kDa CULTURE FILTRATE glycoprotein.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_taxid=1773;
RN [1]
RP SEQUENCE.
RX MEDLINE=96196153; PubMed=86263147
RA Dobos K.M., Khoo K.H., SwiderEK K.M., Brennan P.J., Belisle J.T.;
RT "Definition of the full extent of glycosylation of the 45-kilodalton
RT glycoprotein of Mycobacterium tuberculosis."
RL J. Bacteriol. 178:2498-2506(1996).
SQ SEQUENCE 286 AA; 28764 MW; 4BD7AD89FEB42493 CRC64;

Query Match	89.2%	Score 1539;	DB 2;	Length 286;
Best Local Similarity	99.7%	Pred No. 1.3e-89;		
Matches 285;	Conservative	17 Mismatches	0;	Indels 0;
Gaps	0;			
Qy	40	DPEPAPPVPTTAASPSSTAAAPAPATPVAPPAPAAANTPNAQPGDPNAAAPPADNAPP	99	
Db	1	DPEPAPPVPTTAASPSSTAAAPAPATPVAPPAPAAANTPNAQPGDPNAAAPPADNAPP	60	
Qy	100	PPVIAPNAPQVPRIDNPVGFGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGPPPV	159	
Db	61	PPVIAPNAPQVPRIDNPVGFGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGPPPV	120	
Qy	160	ANDTRIVLGRDLQKLYASAEATDSKAAALGSDMGFFYMPYPGTRINQETVSLDANGVSG	219	
Db	121	ANDTRIVLGRDLQKLYASAEATDSKAAALGSDMGFFYMPYPGTRINQETVSLDANGVSG	180	
Qy	220	SASVYEVKTSKPKNGQIWTGVIGSPAANAPDAGPPQRFVFWLGTANNPVDKGAAL	279	


```

RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0439B06.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AP002882; BAB39873.1; -.
DR Gramene; Q9ASK4; -.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004648; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 698 AA; 72733 MW; F3866E46DCA9DAE6 CRC64;

Query Match 14.4%; Score 248.5; DB 10; Length 698;
Best Local Similarity 30.7%; Pred.No. 7.5e-08;
Matches 89; Conservative 20; Mismatches 134; Indels 47; Gaps 9;

Qy 32 AVPATANADPEAPPPVPTTAASPPSTAAAPPATPVAPPVPPAAATPNAPQGDPAAPP 91
Db 12 ATPATAAASTATPTP--ATQENATPADPSITPPAASPLPSAATTPQPDSPFPPSP 68

Qy 92 PAPPNAPPPVLIAPNAPQVRIQNVGGFSGFALPAGWVESDAHFYGSALLSKTTGDDP 151
Db 69 PAPPVAVPPPATVPPPPVPPVPPVPPPT-----PSATLPPSPSPASVPVPTPATPPKP 121

Qy 152 FPGQPPPVANDTRIVLGRLOKLYASAEATDSKAAARLGSDMGMEFYM-PVPGTGRINQETV 210
Db 122 SPVQQPPVA-----ASPPSSPADLPPPPPPAR--SDTPPVVQSPPPPHRRSRPTP 170

Qy 211 SLDPANGVSGSAYVEKFSKDPKXNGQITGWIGSGPAANPADAGPPQRWFVWLGTTANPP 270
Db 171 WAPPMAPSPGS-----PTKP-----SPASPPIAGDP-----IITPPNP 205

Qy 271 VDRGAAKALAESIRPLVAPPAPAPAEAPAPAPAPAGEVAPTPTTTPQ 320
Db 207 SSPLATPSACSGTGPVVT-ESAPVSGPPSPGTAPATAADRNSKLSLSENTQ 255

RESULT 6
Q7T5D9 PRELIMINARY; PRT; 3288 AA.
AC Q7T5D9,
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

OS	Smian herpes B virus (Cercopithecid herpesvirus 1) (Shbv).
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC	Alphaherpesvirinae; Simplexvirus.
NCBI_TaxID=10325;	
[1]	
RP	SEQUENCE FROM N.A.
RC	STRAIN=E2490;
RX	MEDLINE=26288476; PubMed=12743273;
RA	Perelygina L., Zhu L., Zurkuhlen H., Mills R., Borodovsky M.,
RA	Hilliard J.K.;
RT	"Complete Sequence and Comparative Analysis of Herpes B
RT	virus (Cercopithecine Herpesvirus 1) from a Rhesus Monkey.";
RL	J. Virol. 77:6167-6177 (2003).
DR	EMBL; AF533768; AAP41454.1; -;
SQ	SEQUENCE 3288 AA; 342496 MW; 79C391EE856F7983 CRC64;


```

RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=2248498; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005936; BAC45786.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR006665; OmpA/MotB.
DR InterPro; IPR003882; Pistil extensin.
DR InterPro; IPR000719; Prot_Kinase.
DR Pfam; PF00691; OmpA; 1.
DR PRINTS; PR01218; PS1EXTENSIN.
DR ProDom; PD000001; Prot_Kinase; 1.
DR Complete proteome.
SQ SEQUENCE 745 AA; 74545 MW; 155EDFC74BDF6D CRC64;

Query Match 14.0%; Score 241.5; DB 16; Length 745;
Best Local Similarity 29.6%; Pred. No. 2.2e-07;
Matches 112; Conservative 16; Mismatches 136; Indels 115; Gaps 21;

QY 15 LAALAIAN-----ASALVTAVPATANADPEP-----APP----- 46
DQ 9 LATTALTANQLANTASHAQGASPLVVAQAQOETGPDGKQPPKPPGGAAPAAPARPA 68
QY 47 VPTTAASPSATAAPP---APATVAPPPAAANTNNAQGDPAAPADPNAP-PPV 102
DQ 69 APPAAAPHPHPPAPPAAPRAAPPPPP-PPAARAPPPPPPPPPPAKQSPPPA 126
QY 103 IAP--NAPQVRIIDNVGSGFALLPAGWVESDAHFYGSALLSKTTGDP--PFGQPPV 159
DQ 127 AAPQQAFTP---PPAPPAARPAPTPPAPPAAPQAHAPPPPPPPAARPTPTPPPPPA 183
QY 160 ANDRIVLGRDQKLYASAEATDSKAA-----ARLGSMDGEFMYEYG-----TRI 205
DQ 184 GPAARPT-----PAPTATPTFVAPPAAPTARPGS-----PAPATPATPTTPA 227
QY 206 NQETVSLDANGVSGSASYEYKFSKPPKPNQIWTGVIGSPAANAPAGPPQRFVVMWL 265
DQ 228 PTATPATPATAPGSTP-----GAPPAGPGAPPG-----VRPG 262
QY 266 T---ANNPVDKGAALAESTRP--LVAPP-----PA--PAPAPAEAP--APAPAGEVA 311
DQ 263 SPPAAGSPAPGATPATTTTPAGGTATPFGSGRGPASTPAPGATPATPATPAPGGALT 322
QY 312 PTP-----TPTPTORTLP 324
DQ 323 PPPGPGAGPTFGPGGTP 341

RESULT 10
Q7T591
ID Q7T591 PRELIMINARY; PRT; 3326 AA.
AC Q7T591;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Large tegument protein.
GN UL36.
OS Simian herpes B virus (Cercopithecoid herpesvirus 1) (Shbv).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10325;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22607624; PubMed=12721804;

RA Ohsawa K., Black D.H., Sato H., Rogers K., Eberle R.;
RT "Sequence and genetic arrangement of the UL region of the monkey B
RT virus (Cercopithecine herpesvirus 1) genome and comparison with the UL
RT region of other primate herpesviruses.";
RL Arch. Virol. 148:989-997(2003).
DR EMBL; AB096160; BAC59076.2; -.
SQ SEQUENCE 3326 AA; 345566 MW; 6B53E3860F43CDF0 CRC64;

Query Match 13.9%; Score 240.5; DB 12; Length 3326;
Best Local Similarity 29.1%; Pred. No. 1.3e-06;
Matches 93; Conservative 24; Mismatches 162; Indels 41; Gaps 9;

QY 17 ALAIAWASASLVTVAVPATANADPEAPVPPTTAAASPSSTAAAPAPATPVAP--PPPA 74
DQ 2875 AAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAPAA 2934
QY 75 AANTPNAQCDPNAPPPADPNAPPPV--IAPNAPQVRIIDNVGSGFALLPAGWVESD 132
DQ 2935 AVPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPAAPATAT 2994
QY 133 AAHFDYGSALLSKTTGDPFFGQP-----PPVANDTRIVLG-----RLDQKLYASAEATD 182
DQ 2995 APAATPASFPVPTTSSLPTTSPKPPAFFQFSLATGSGVAPGDFRRRAPSRPTAAVPAAP 3054
QY 183 SKAAAR-----LGSDMGEFYMPYPGTRINQETVSLDANGVSGSASYEYKFSKPPNG 236
DQ 3055 SRPPARLARPAVSRSTESFALP-PDELARPTPEAPAPTETEAPVAERPAPEPPQG 3113
QY 237 QIWTGVIGSPAANAPAGPPQRFVVMWLGTANNPVDKGAALAEISIRLVAPP-----PA 292
DQ 3114 R-----PPSPAAPDAGP-----AAASGSGVGPAPRLGALVPGRVAPRRQIPP 3157
QY 293 PAPAPAEAPAPAPAGEVAP 312
DQ 3158 PAPPREIPAPPPPPRSHAP 3177

RESULT 11
Q8W5K6
ID Q8W5K6 PRELIMINARY; PRT; 1269 AA.
AC Q8W5K6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSJNBA0079B05.10 OR OSJNAA0079B05.2.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Sasaki C., Henry D., Oates R., Simmons J.;
RT "Rice Genomic Sequence.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Yang T.J., Nah G., Soderlund C., Chen M., Kim H.-R.,
RA Rambo T., Sasaki C., Henry D., Oates R., Simmons J.;
RT "Rice Genomic Sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079179; AAL31655.1; -.
DR EMBL; AC116601; AAM08709.1; -.
DR Gramene; Q8W5K6; -.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR008973; C2.CaLB.
DR InterPro; IPR003104; FH2.
DR Pfam; PF02181; FH2; 1.
DR SMART; SM00498; FH2; 1.
DR Hypothetical protein.
SQ SEQUENCE 1269 AA; 138432 MW; 262E546481B25CA6 CRC64;

```

Query Match	13.8%; Score 238; DB 10; Length 1269;
Best Local Similarity	27.2%; Pred.No. 6.6e-07;
Matches	82; Conservative 20; Mismatches 121; Indels 78; Gaps 10;
QY	36 TANADPECAPPPVPTTAASGSPSTAAPAPATPVAPPAAAAAATFNAQGDPNAAAPP-----91
Db	540 TAAAPPPPPPPPPPSGNKP--AFSPPPPPPPPPPLPQSNIYASSQPPIPPPPPLENC597
QY	92 --PADNAPPPPVIAPN--AQPVRIINPVGGFSFALPAGWVESDAAHFDYSALLSKIT147
Db	598 LVPSEPPPPPPPPPLPNRSVPPPPPPPPPLFNHSL-----633
QY	148 GDPFPFGOPPPVANTRIVLGRLOKLYASAEATDSKAAARLGSDMGFEFMPYPVGTRINQ207
Db	634 -PPPPPPPPPSLFNLV-----PPPAFCIGNKFPAPPPPPPPPRSSS676
QY	208 ETVISLDANGVSGSAYEVEKSDSKNGKIWTGVIGSPAANAPDAGPQRWFVVWLGTAT267
Db	677 RTPTGAATSSKGPPPPPPPPPLPFANRNP-----GVPSAPPPPPPPP-----ANRS724
QY	268 NNPVDKGAAKALAESIRFLVAPPPAPAPA---PAPAPAPAPAGAVATPTTTPORTLT323
Db	725 NGP-----SAPAPL--PPPLPAAANKRPAPPPPPPLMTGKAPAPPPPPQAPK773
QY	324 P 324
Db	774 P 774
RESULT 12	
Q7XH56	PRELIMINARY; PRT; 1269 AA.
ID Q7XH56	AC Q7XH56;
DT 01-OCT-2003	(TrEMBLrel. 25, Created)
DT 01-OCT-2003	(TrEMBLrel. 25, Last sequence update)
DE 01-OCT-2003	(TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.	
CN OSUNBA0079805.10.	
OS Oryza sativa [japonica cultivar-group].	
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
OC Ehrhartoideae; OrYZeae; Oryza.	
OX NCBI_TaxID=39947;	
[1] RN	SEQUENCE FROM N.A.
RP STRAIN=cv. Nipponbare;	
RC The Rice Chromosome 10 Sequencing Consortium;	
RA "In-depth view of structure, activity, and evolution of rice	
RT chromosome 10.";	
RN Science 300:1566-1569(2003).	
[2]	
RP SEQUENCE FROM N.A.	
RC STRAIN=cv. Nipponbare;	
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;	
RL Submitted (MAY-2003) to the ENBL/GenBank/DBJ databases.	
DR EMBL; AE017051; AAP51898.1;	
KL Hypothetical protein.	
SQ SEQUENCE 1269 AA; 138431 MW; 262E546481B25CA6 CRC64;	

Db	634	--PPPPPPPPPSLNRLV-----PPPPAGIGNKFPAPPPPPPPPPSS 676
Qy	208	ETVSLDANGVSGSASYEVKFSDFSKPNQIGTVIGSPAANAPDAGPPQRFVFWLGT 267
Db	677	RTPTGAATSKGPPPPPPPLPANTNGP-----GVPSADPPPPPPPP-----ANRS 724
Qy	268	NNFVDKGAALKALAESIRPLVAPPAPAPA-----PAPAPAPAGVAPPTTPTPTQRTL 323
Db	725	NGP-----SAPAPPL--PPPLFAAANKRNPPAPPPPLMTGKAPAPPPPPQAPK 773
Qy	324	P 324
Db	774	P 774
RESULT 13		
ID	Q41805	
ID	Q41805	PRELIMINARY; PRT; 1188 AA.
AC	Q41805;	
DT	01-NOV-1996	(T-EMBLrel. 01, Created)
DT	01-NOV-1996	(T-EMBLrel. 01, Last sequence update)
DT	01-OCT-2003	(T-EMBLrel. 25, Last annotation update)
DE	Extensin-like protein precursor.	
OS	Zea mays (Maize);	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
OC	PACCAD clade; Panicoideae; Andropogoneae; Zea.	
OX	NCBI_TaxID=4577;	
ON	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=B73; TISSUE=Pollen;	
RA	Rubinstein A.L.; Broadwater A.H., Lowrey K., Bedinger P.A.;	
RL	"Pex genes: pollen-specific genes with extensin-like domains.";	
RT	Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.	
DR	EMBL;	Z34465; CAA84230.1; -.
DR	PIR;	S49915; S49915.
DR	InterPro;	IPR001611; LRR.
DR	InterPro;	IPR007090; LRR_plant.
DR	Pfam;	PF00560; LRR; 3.
KW	Signal.	
FT	SIGNAL	
SEQ	SEQUENCE	1 27 POTENTIAL.
		1188 AA; 120981 MW; 2C77CTFBD7130149 CRC64;
Query Match		
Best Local Similarity 28.9%; Score 237.5; DE 10; Length 1188;		
Matches 87; Conservative 30; Mismatches 123; Indels 61; Gaps 13;		
Qy	34	PATANADBP--APPVTTAAAGPSTAAAPPATPVAPPPPPAAANTENAQGDPAAPP 91
Db	569	PAPVAGPPPPVKSPPPPTLVASPPPPVKSPPPPA-PVASPPPPVKSP---PTPVASPP 624
Qy	92	PADPNA-PPPVVIAPNAQPFVRIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDP 150
Db	625	PPAPVASSPPPMKSP--PPPTPVSSPPPPPEKSPPPPPPAKSTPPPEEYTPPTTSVKSSPP 682
Qy	151	PPFGQPPPVANDRIVLRDLQKLYASATQSKAAARLGSMDGGEFVMPYVPGTRINQETV 210
Db	683	PEKSLPP-----TLIPSPPOEK--PTPSTTSKPPS-----SPEKP 718
Qy	211	SLDANGVSGSASYEVKFSDFSKPNQIGTVIGSPAANAPDAGPPQRFVFWLGTANNP 270
Db	719	SPKPEVPSPQPQ--TPKSPPPAPVSSPPPTPVSSPPALAPVSSPP-----SVKSSP 768
Qy	271	VDKGAALKALAESIRPLVAPPAP-----APAPAE---PAPAPAGVAPPTTPTPTQRTL 323
Db	769	PP-----APLSSPPAPQVKSSPPPVQVSSFPAPKSPPLAPVSSPPQVEKTS 817
Qy	324	P 324
Db	818	P 818

RESULT 15	
Q82DU7	
ID Q82DU7	PRELIMINARY;
	PRT; 584 AA.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 18:21:10 ; Search time 16.5 Seconds
(without alignments)
1016.875 Million cell updates/sec

Title: US-10-720-192-3
Perfect score: 1726
Sequence: 1 MHQVDNLTERRKGRLAALAI.....PAGEVAPTTPTPTORTLPA 325

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/FACTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1726	100.0	325	1	US-08-382-184-2
2	1726	100.0	325	2	US-08-641-356-2
3	1726	100.0	325	3	US-09-132-528-2
4	1726	100.0	325	3	US-09-132-528-3
5	1726	100.0	325	3	US-08-875-494-2
6	1726	100.0	325	4	US-09-599-366-2
7	1726	100.0	325	4	US-09-599-366-3
8	1726	100.0	325	4	US-08-875-494-2
9	1726	100.0	325	4	US-09-510-031A-6
10	1726	100.0	325	4	US-09-985-372-2
11	1726	100.0	325	4	US-09-985-372-3
12	1726	100.0	332	3	US-08-818-111-53
13	1726	100.0	332	4	US-08-818-111-53
14	1726	100.0	332	4	US-09-056-556-53
15	1726	100.0	332	4	US-09-072-596-53
16	1726	100.0	332	4	US-09-072-967-53
17	1720	99.7	325	4	US-09-510-031A-2
18	1542	89.3	286	1	US-08-382-184-3
19	1542	89.3	286	2	US-08-641-356-3
20	1542	89.3	286	3	US-09-132-528-4
21	1542	89.3	286	3	US-08-875-494-3
22	1542	89.3	286	4	US-09-599-366-4
23	1542	89.3	286	4	US-08-875-494-3
24	1542	89.3	286	4	US-09-985-372-4
25	1539	89.2	286	4	US-09-510-031A-7
26	1499	86.8	652	4	US-09-072-596-350
27	1499	86.8	652	4	US-09-072-967-355

28	1499	86.8	802	4	US-09-056-556-214	Sequence 214, App
29	1499	86.8	802	4	US-09-072-596-209	Sequence 209, App
30	1499	86.8	802	4	US-09-072-596-346	Sequence 346, App
31	1499	86.8	802	4	US-09-072-967-214	Sequence 214, App
32	1499	86.8	802	4	US-09-072-967-351	Sequence 351, App
33	1499	86.8	802	4	US-09-287-849-10	Sequence 10, Appl
34	985	57.1	381	4	US-09-510-031A-5	Sequence 5, Appl
35	980	56.8	287	4	US-08-311-731A-96	Sequence 96, Appl
36	980	56.8	287	4	US-09-510-031A-4	Sequence 4, Appl
37	969	56.1	287	4	US-09-510-031A-8	Sequence 8, Appl
38	233	13.5	8991	4	US-08-714-741-32	Sequence 32, Appl
39	200.5	11.6	214	1	US-08-217-327-4	Sequence 4, Appl
40	200	11.6	408	1	US-07-609-716-65	Sequence 65, Appl
41	200	11.6	408	3	US-08-475-411A-65	Sequence 65, Appl
42	200	11.6	408	3	US-08-478-029A-65	Sequence 65, Appl
43	197.5	11.4	330	1	US-08-642-255-32	Sequence 32, Appl
44	197	11.4	538	4	US-09-616-289-43	Sequence 43, Appl
45	196.5	11.4	666	4	US-09-050-739-70	Sequence 70, Appl

ALIGNMENTS

RESULT 1

US-08-382-184-2

; Sequence 2, Application US/08382184

; Patent No. 5714593

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: MICROBACTERIAL PROTEINS,

; TITLE OF INVENTION: MICROORGANISMS PRODUCING THEM AND THEIR USE FOR VACCINES
AND FOR THE DETECTION OF TUBERCULOSIS

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER &

; STREET: 1755 S. Jefferson Davis Highway, Suite 400

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/382,184

; FILING DATE: 01-FEB-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: OBLON, NORMAN F

; REGISTRATION NUMBER: 24,618

; REFERENCE/DOCKET NUMBER: 22640720

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-413-3000

; TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 325 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-382-184-2

Query Match 100.0%; Score 1726; DB 1; Length 325;

Best Local Similarity 100.0%; Pred No. 6.4e-127;

Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 MHQVDNLTERRKGRLAALAIAMASASLVTVAVPATANADPEPFPVPTTAASPESTAAA 60

Db

1 MHQVDNLTERRKGRLAALAIAMASASLVTVAVPATANADPEPFPVPTTAASPESTAAA 60

QY 61 PPAPATPVAPPPAAAANTPNAQFGDPNAAAPPADPNAPPPVIAFNAPQPVRIIDNPVGGF 120
 Db 61 PPAPATPVAPPPAAAANTPNAQFGDPNAAAPPADPNAPPPVIAFNAPQPVRIIDNPVGGF 120
 QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGPPGPPVANDTRIVLGRDOKLYASAEA 180
 Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGPPGPPVANDTRIVLGRDOKLYASAEA 180
 QY 181 TDSKAAARLGSMDGFEFYPYPGTRINQETVSLDANGVSGSASYEYKFSDFSKPQGIWT 240
 Db 181 TDSKAAARLGSMDGFEFYPYPGTRINQETVSLDANGVSGSASYEYKFSDFSKPQGIWT 240
 QY 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVDKGAALKALAESIRPLVAPPAPAPAPAP 300
 Db 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVDKGAALKALAESIRPLVAPPAPAPAPAP 300
 QY 301 APAPAPAGEVAPTPTTPTQRTPLA 325
 Db 301 APAPAPAGEVAPTPTTPTQRTPLA 325

RESULT 2
 US-08-641-356-2
 ; Sequence 2, Application US/08641356
 ; Patent No. 5866130
 ; GENERAL INFORMATION:
 ; APPLICANT: MICROBACTERIAL PROTEINS,
 ; TITLE OF INVENTION: MICROORGANISMS PRODUCING THEM AND THEIR USE FOR VACCINES
 ; TITLE OF INVENTION: AND FOR THE DETECTION OF TUBERCULOSIS
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER &
 ; ADDRESS: NEUSTADT, P.C.
 ; STREET: 1755 S. Jefferson Davis Highway, Suite 400
 ; CITY: Arlington
 ; STATE: Virginia
 ; COUNTRY: U.S.A.
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/641,356
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/382,184
 ; FILING DATE: 01-FEB-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 22640720
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-413-3000
 ; TELEFAX: 703-413-2220
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 325 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-641-356-2

Query Match 100.0%; Score 1726; DB 2; Length 325;
 Best Local Similarity 100.0%; Pred. No. 6.4e-127;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 PPAPATPVAPPPAAAANTPNAQFGDPNAAAPPADPNAPPPVIAFNAPQPVRIIDNPVGGF 120
 Db 61 PPAPATPVAPPPAAAANTPNAQFGDPNAAAPPADPNAPPPVIAFNAPQPVRIIDNPVGGF 120
 QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGPPGPPVANDTRIVLGRDOKLYASAEA 180
 Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGPPGPPVANDTRIVLGRDOKLYASAEA 180
 QY 181 TDSKAAARLGSMDGFEFYPYPGTRINQETVSLDANGVSGSASYEYKFSDFSKPQGIWT 240
 Db 181 TDSKAAARLGSMDGFEFYPYPGTRINQETVSLDANGVSGSASYEYKFSDFSKPQGIWT 240
 QY 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVDKGAALKALAESIRPLVAPPAPAPAPAP 300
 Db 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVDKGAALKALAESIRPLVAPPAPAPAPAP 300
 QY 301 APAPAPAGEVAPTPTTPTQRTPLA 325
 Db 301 APAPAPAGEVAPTPTTPTQRTPLA 325

RESULT 3
 US-09-132-528-2
 ; Sequence 2, Application US/09132528A
 ; Patent No. 6221353
 ; GENERAL INFORMATION:
 ; APPLICANT: LAQUEYREIRB, Anne
 ; APPLICANT: MARCHAL, Gilles
 ; APPLICANT: PESCHER, Pascale
 ; APPLICANT: ROWAIN, Felix
 ; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICORORGANISMS PRODUCING THEM
 ; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
 ; TITLE OF INVENTION: TUBERCULOSIS
 ; FILE REFERENCE: 0660-0138-ODIV
 ; CURRENT APPLICATION NUMBER: US/09/132,528A
 ; CURRENT FILING DATE: 1998-08-11
 ; EARLIER APPLICATION NUMBER: 08/641,356
 ; EARLIER FILING DATE: 1996-04-30
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 325
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 ; US-09-132-528-2

Query Match 100.0%; Score 1726; DB 3; Length 325;
 Best Local Similarity 100.0%; Pred. No. 6.4e-127;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHQVDNLTFRKGRLLAALAIAAASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
 Db 1 MHQVDNLTFRKGRLLAALAIAAASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
 QY 61 PPAPATPVAPPPAAAANTPNAQFGDPNAAAPPADPNAPPPVIAFNAPQPVRIIDNPVGGF 120
 Db 61 PPAPATPVAPPPAAAANTPNAQFGDPNAAAPPADPNAPPPVIAFNAPQPVRIIDNPVGGF 120
 QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGPPGPPVANDTRIVLGRDOKLYASAEA 180
 Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFGPPGPPVANDTRIVLGRDOKLYASAEA 180
 QY 181 TDSKAAARLGSMDGFEFYPYPGTRINQETVSLDANGVSGSASYEYKFSDFSKPQGIWT 240
 Db 181 TDSKAAARLGSMDGFEFYPYPGTRINQETVSLDANGVSGSASYEYKFSDFSKPQGIWT 240
 QY 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVDKGAALKALAESIRPLVAPPAPAPAPAP 300
 Db 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVDKGAALKALAESIRPLVAPPAPAPAPAP 300
 QY 301 APAPAPAGEVAPTPTTPTQRTPLA 325

```
Db 301 APAPAPAGEVAPTPTTPTPQRTLPA 325
RESULT 4
US-09-132-528-3
; Sequence 3, Application US/09132528A
; Patent No. 6221353
; GENERAL INFORMATION:
; APPLICANT: LAQUEYRERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/132,528A
; EARLIER FILING DATE: 1998-08-11
; EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-132-528-3

Query Match 100.0%; Score 1726; DB 3; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDNLTTRKGRGLAALAIAAMASASLVTVAVPATANADPPAPVPVPTTAASPPSTAAA 60
Db 1 MHQVDNLTTRKGRGLAALAIAAMASASLVTVAVPATANADPPAPVPVPTTAASPPSTAAA 60
Qy 61 PPAPATPVAPPPPPAAANTPNAQPDNAPPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
Db 61 PPAPATPVAPPPPPAAANTPNAQPDNAPPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPPVANDTRIVLGRDQKLYASAEA 180
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPPVANDTRIVLGRDQKLYASAEA 180
Qy 181 TDSKAAARLGSDMGFEYMPYGTTRINQETVSLDANGVSGSASYEVKFSDPKPNQIWT 240
Db 181 TDSKAAARLGSDMGFEYMPYGTTRINQETVSLDANGVSGSASYEVKFSDPKPNQIWT 240
Qy 241 GVIGSPAANADAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300
Db 241 GVIGSPAANADAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300
Qy 301 APAPAPAGEVAPTPTTPTPQRTLPA 325
Db 301 APAPAPAGEVAPTPTTPTPQRTLPA 325

RESULT 5
US-08-875-494-2
; Sequence 2, Application US/08875494
; Patent No. 6221610
; GENERAL INFORMATION:
; APPLICANT: LAQUEYRERIE, ANNE
; APPLICANT: MARCHAL, GILLES
; APPLICANT: PESCHER, PASCALE
; APPLICANT: ROMAIN, FELIX
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USES FOR VACCINES AND FOR THE DETECTION OF
; FILE REFERENCE: 0660-0122-0 PCT
; CURRENT APPLICATION NUMBER: US/08/875,494
; EARLIER FILING DATE: 1997-08-01
; EARLIER APPLICATION NUMBER: PCT/FR96/00165

Db 301 APAPAPAGEVAPTPTTPTPQRTLPA 325
US-09-132-528-3

Query Match 100.0%; Score 1726; DB 3; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDNLTTRKGRGLAALAIAAMASASLVTVAVPATANADPPAPVPVPTTAASPPSTAAA 60
Db 1 MHQVDNLTTRKGRGLAALAIAAMASASLVTVAVPATANADPPAPVPVPTTAASPPSTAAA 60
Qy 61 PPAPATPVAPPPPPAAANTPNAQPDNAPPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
Db 61 PPAPATPVAPPPPPAAANTPNAQPDNAPPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPPVANDTRIVLGRDQKLYASAEA 180
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPPVANDTRIVLGRDQKLYASAEA 180
Qy 181 TDSKAAARLGSDMGFEYMPYGTTRINQETVSLDANGVSGSASYEVKFSDPKPNQIWT 240
Db 181 TDSKAAARLGSDMGFEYMPYGTTRINQETVSLDANGVSGSASYEVKFSDPKPNQIWT 240
Qy 241 GVIGSPAANADAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300
Db 241 GVIGSPAANADAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300
Qy 301 APAPAPAGEVAPTPTTPTPQRTLPA 325
Db 301 APAPAPAGEVAPTPTTPTPQRTLPA 325

RESULT 6
US-09-599-366-2
; Sequence 2, Application US/09599366
; Patent No. 6335181
; GENERAL INFORMATION:
; APPLICANT: LAQUEYRERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/599,366
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 09/132,528
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 08/641,356
; PRIOR FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-599-366-2

Query Match 100.0%; Score 1726; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDNLTTRKGRGLAALAIAAMASASLVTVAVPATANADPPAPVPVPTTAASPPSTAAA 60
Db 1 MHQVDNLTTRKGRGLAALAIAAMASASLVTVAVPATANADPPAPVPVPTTAASPPSTAAA 60
```

Db 1 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
Qy 61 PPAPATPVAPPPPPAAANTENACPDENAAAPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120
Db 61 PPAPATPVAPPPPPAAANTENACPDENAAAPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPGVVANDTRIVLGRDOKLYASAEA 180
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPGVVANDTRIVLGRDOKLYASAEA 180
Qy 181 TDSKAAARLGSDMGGEFYMYPGTRINQETVSLDANGVSGSASYEVKFSDPKNGQIWT 240
Db 181 TDSKAAARLGSDMGGEFYMYPGTRINQETVSLDANGVSGSASYEVKFSDPKNGQIWT 240
Qy 241 GVIGSPAANAPDAGPPQRFVWLVGTANNPVDKGAAKALAESIRPLVAPPAPAPAPAP 300
Db 241 GVIGSPAANAPDAGPPQRFVWLVGTANNPVDKGAAKALAESIRPLVAPPAPAPAPAP 300
Qy 301 APAPAGEVAPTPTTPTPQRTPLA 325
Db 301 APAPAGEVAPTPTTPTPQRTPLA 325

RESULT 7

US-09-599-366-3

; Sequence 3, Application US/09599366

; Patent No. 6335181

; GENERAL INFORMATION:

; APPLICANT: LAQUEYRIERIE, Anne

; APPLICANT: MARCHAL, Gilles

; APPLICANT: PESCHER, Pascale

; APPLICANT: ROMAIN, Felix

; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM

; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF

; FILE REFERENCE: 0660-0138-ODIV

; CURRENT APPLICATION NUMBER: US/09/599,366

; PRIOR FILING DATE: 2000-06-21

; PRIOR FILING DATE: 1998-08-11

; PRIOR FILING DATE: 1998-04-30

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 325

; TYPE: PRT

; ORGANISM: Mycobacterium tuberculosis

US-09-599-366-3

Query Match 100.0%; Score 1726; DB 4; Length 325;

Best Local Similarity 100.0%; Pred. No. 6.4e-127;

Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60

Db 1 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60

Qy 61 PPAPATPVAPPPPPAAANTENACPDENAAAPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120

Db 61 PPAPATPVAPPPPPAAANTENACPDENAAAPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120

Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPGVVANDTRIVLGRDOKLYASAEA 180

Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPGVVANDTRIVLGRDOKLYASAEA 180

Qy 181 TDSKAAARLGSDMGGEFYMYPGTRINQETVSLDANGVSGSASYEVKFSDPKNGQIWT 240

Db 181 TDSKAAARLGSDMGGEFYMYPGTRINQETVSLDANGVSGSASYEVKFSDPKNGQIWT 240

Qy 241 GVIGSPAANAPDAGPPQRFVWLVGTANNPVDKGAAKALAESIRPLVAPPAPAPAPAP 300

Db 241 GVIGSPAANAPDAGPPQRFVWLVGTANNPVDKGAAKALAESIRPLVAPPAPAPAPAP 300

Qy 301 APAPAGEVAPTPTTPTPQRTPLA 325
Db 301 APAPAGEVAPTPTTPTPQRTPLA 325

RESULT 8

US-08-875-494-2

; Sequence 2, Application US/08875494

; Patent No. 6379902

; GENERAL INFORMATION:

; APPLICANT: LAQUEYRIERIE, ANNE

; APPLICANT: MARCHAL, GILLES

; APPLICANT: PESCHER, PASCALE

; APPLICANT: ROMAIN, FELIX

; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM

; TITLE OF INVENTION: AND THEIR USES FOR VACCINES AND FOR THE DETECTION OF

; FILE REFERENCE: 0660-0122-0 PCT

; CURRENT APPLICATION NUMBER: US/08/875,494

; CURRENT FILING DATE: 1997-08-01

; PRIOR APPLICATION NUMBER: PCT/FR96/00166

; PRIOR FILING DATE: 1996-02-01

; PRIOR APPLICATION NUMBER: 382184

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 325

; TYPE: PRT

; ORGANISM: Mycobacterium tuberculosis

US-08-875-494-2

Query Match 100.0%; Score 1726; DB 4; Length 325;

Best Local Similarity 100.0%; Pred. No. 6.4e-127;

Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60

Db 1 MHQVDPNLTTRKGRLLAALAIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60

Qy 61 PPAPATPVAPPPPPAAANTENACPDENAAAPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120

Db 61 PPAPATPVAPPPPPAAANTENACPDENAAAPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120

Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPGVVANDTRIVLGRDOKLYASAEA 180

Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPDPGPPGVVANDTRIVLGRDOKLYASAEA 180

Qy 181 TDSKAAARLGSDMGGEFYMYPGTRINQETVSLDANGVSGSASYEVKFSDPKNGQIWT 240

Db 181 TDSKAAARLGSDMGGEFYMYPGTRINQETVSLDANGVSGSASYEVKFSDPKNGQIWT 240

Qy 241 GVIGSPAANAPDAGPPQRFVWLVGTANNPVDKGAAKALAESIRPLVAPPAPAPAPAP 300

Db 241 GVIGSPAANAPDAGPPQRFVWLVGTANNPVDKGAAKALAESIRPLVAPPAPAPAPAP 300

Qy 301 APAPAGEVAPTPTTPTPQRTPLA 325

Db 301 APAPAGEVAPTPTTPTPQRTPLA 325

RESULT 9

US-09-510-031A-6

; Sequence 6, Application US/09510031A

; Patent No. 6638518

; GENERAL INFORMATION:

; APPLICANT: Ratliff, Timothy

; APPLICANT: Kline, Joel

; TITLE OF INVENTION: METHOD FOR INHIBITING INFLAMMATORY RESPONSES

; FILE REFERENCE: 140.0010.0101

; CURRENT APPLICATION NUMBER: US/09/510,031A

; CURRENT FILING DATE: 2000-02-22

```
; PRIOR APPLICATION NUMBER: US 60/121,177
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 6
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-510-031A-6

Query Match      100.0%; Score 1726; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDNLTTRKGLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
Db 1 MHQVDNLTTRKGLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
Qy 61 PPAPATPVAPPPPPAAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 120
Db 61 PPAPATPVAPPPPPAAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 120
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQGPVANDTRIVLGRDOKLYASAEA 180
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQGPVANDTRIVLGRDOKLYASAEA 180
Qy 181 TDSKAAARLGSMDGFEFMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 240
Db 181 TDSKAAARLGSMDGFEFMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 240
Qy 241 GVIGSPAANADAGPPQRFVWVLTANNPVDKGAALAEIRPLVAPPAPAPAPAP 300
Db 241 GVIGSPAANADAGPPQRFVWVLTANNPVDKGAALAEIRPLVAPPAPAPAPAP 300
Qy 301 APAPAGEVAPTPTTPTPQRTLPA 325
Db 301 APAPAGEVAPTPTTPTPQRTLPA 325

RESULT 11
US-09-985-372-3
; Sequence 3, Application US/09985372
; Patent No. 6676945
; GENERAL INFORMATION:
; APPLICANT: LAQUEYERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/985,372
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 09/132,528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-985-372-3

Query Match      100.0%; Score 1726; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDNLTTRKGLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
Db 1 MHQVDNLTTRKGLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
Qy 61 PPAPATPVAPPPPPAAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 120
Db 61 PPAPATPVAPPPPPAAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 120
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQGPVANDTRIVLGRDOKLYASAEA 180
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPQGPVANDTRIVLGRDOKLYASAEA 180
Qy 181 TDSKAAARLGSMDGFEFMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 240
Db 181 TDSKAAARLGSMDGFEFMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 240
Qy 241 GVIGSPAANADAGPPQRFVWVLTANNPVDKGAALAEIRPLVAPPAPAPAPAP 300
Db 241 GVIGSPAANADAGPPQRFVWVLTANNPVDKGAALAEIRPLVAPPAPAPAPAP 300
Qy 301 APAPAGEVAPTPTTPTPQRTLPA 325
Db 301 APAPAGEVAPTPTTPTPQRTLPA 325

RESULT 10
US-09-985-372-2
; Sequence 2, Application US/09985372
; Patent No. 6676945
; GENERAL INFORMATION:
; APPLICANT: LAQUEYERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/985,372
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 09/132,528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-985-372-2

Query Match      100.0%; Score 1726; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDNLTTRKGLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
Db 1 MHQVDNLTTRKGLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
Qy 61 PPAPATPVAPPPPPAAAANTPNAQPGDPNAAAPPADPNAPPPVPIAPNAPQVRIIDNPVGGF 120
```

RESULT 12
US-08-818-112-53
; Sequence 53, Application US/08818112
; Patent No. 6290969
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 133
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,112
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-818-112-53

Query Match 100.0%; Score 1726; DB 3; Length 332;
Best Local Similarity 100.0%; Pred. No. 6.5e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHQVDPNLTTRKGRLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
DB 8 MHQVDPNLTTRKGRLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 67
QY 61 PPAPATPVAPPPAAANTENAGDPNAPPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120
DB 68 PPAPATPVAPPPAAANTENAGDPNAPPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 127
QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGOPPPVANDTRIVLGRDQKLYASAEA 180
DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGOPPPVANDTRIVLGRDQKLYASAEA 187
QY 181 TDSKAARLGSDMGCFYMPYPCGTRINQETVSLDANGVSGSASYEVKFSKPKNGQIWT 240
DB 188 TDSKAARLGSDMGCFYMPYPCGTRINQETVSLDANGVSGSASYEVKFSKPKNGQIWT 247
QY 241 GVIGSPAANADAGPPQRFVWVWLGTANNPVDKGAALAESIRPLVAPPPAPAPAPAE 300
DB 248 GVIGSPAANADAGPPQRFVWVWLGTANNPVDKGAALAESIRPLVAPPPAPAPAPAE 307
QY 301 APAPAPAGEVAPTTPTPTQRTPLA 325
|||||

Db 308 APAPAPAGEVAPTTPTPTQRTPLA 332
RESULT 13
US-08-818-111-53
; Sequence 53, Application US/08818111
; Patent No. 6338852
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TUBERCULOSIS
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,111
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-818-111-53

Query Match 100.0%; Score 1726; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 6.5e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHQVDPNLTTRKGRLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
DB 8 MHQVDPNLTTRKGRLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 67
QY 61 PPAPATPVAPPPAAANTENAGDPNAPPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120
DB 68 PPAPATPVAPPPAAANTENAGDPNAPPPADPNAPPPVIAPNAPQPVRIIDNPVGGF 127
QY 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGOPPPVANDTRIVLGRDQKLYASAEA 180
DB 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGOPPPVANDTRIVLGRDQKLYASAEA 187
QY 181 TDSKAARLGSDMGCFYMPYPCGTRINQETVSLDANGVSGSASYEVKFSKPKNGQIWT 240
DB 188 TDSKAARLGSDMGCFYMPYPCGTRINQETVSLDANGVSGSASYEVKFSKPKNGQIWT 247
QY 241 GVIGSPAANADAGPPQRFVWVWLGTANNPVDKGAALAESIRPLVAPPPAPAPAPAE 300
DB 248 GVIGSPAANADAGPPQRFVWVWLGTANNPVDKGAALAESIRPLVAPPPAPAPAPAE 307
QY 301 APAPAPAGEVAPTTPTPTQRTPLA 325
|||||

Db 308 APAPAPAGEVAPTPTTPTPTPTLPA 332
|||||
RESULT 14
US-09-056-556-53
; Sequence 53, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
; NUMBER OF SEQUENCES: 241
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,556
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-056-556-53
Query Match 100.0%; Score 1726; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 6.5e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHQVDNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
Db 8 MHQVDNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 67
Qy 61 PPAPATPVAPPPAAANTPNAQPGDPNAPPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
Db 68 PPAPATPVAPPPAAANTPNAQPGDPNAPPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 127
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPPVANDTRIVLGRLDQKLYASAEA 180
Db 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPPVANDTRIVLGRLDQKLYASAEA 187
Qy 181 TDSKAAARLGSDMGFEYMPYGTTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 240
Db 188 TDSKAAARLGSDMGFEYMPYGTTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 247
Qy 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300
Db 248 GVIGSPAANAPDAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 307
Qy 301 APAPAPAGEVAPTPTTPTPTPTLPA 325
Db 308 APAPAPAGEVAPTPTTPTPTPTLPA 332

RESULT 15
US-09-072-596-53
; Sequence 53, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Wedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-072-596-53

Query Match 100.0%; Score 1726; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 6.5e-127;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHQVDNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 60
Db 8 MHQVDNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA 67
Qy 61 PPAPATPVAPPPAAANTPNAQPGDPNAPPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 120
Db 68 PPAPATPVAPPPAAANTPNAQPGDPNAPPPADPNAPPPVPIAPNAPQPVRIIDNPVGGF 127
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPPVANDTRIVLGRLDQKLYASAEA 180
Db 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPPVANDTRIVLGRLDQKLYASAEA 187
Qy 181 TDSKAAARLGSDMGFEYMPYGTTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 240
Db 188 TDSKAAARLGSDMGFEYMPYGTTRINQETVSLDANGVSGSASYEYVKFSDPSKNGQIWT 247
Qy 241 GVIGSPAANAPDAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300
Db 248 GVIGSPAANAPDAGPPQRFVWVLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 307
Qy 301 APAPAPAGEVAPTPTTPTPTPTLPA 325

Thu Jul 8 11:11:31 2004

us-10-720-192-3.ra1

Page 8

Db 308 APAPAPAGEVAPTPTTPTQRTLPA 332
|||||

Search completed: July 7, 2004, 18:24:53
Job time : 17.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 18:24:06 ; Search time 42.5 Seconds

(without alignments)
2380.406 Million cell updates/sec

Title: US-10-720-192-3

Perfect score: 1726

Sequence: 1 MHQVDNLTERRKGRLLAALAI.....PAGEVAPPTTPTPQRTLPA 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query *	Score	Match	Length	DB	ID	Description
1	1726	100.0	325	10	US-09-985-372-2		Sequence 2, Appli
2	1726	100.0	325	10	US-09-985-372-3		Sequence 3, Appli
3	1726	100.0	332	12	US-09-886-349A-41		Sequence 41, Appl
4	1726	100.0	332	14	US-10-193-002-53		Sequence 53, Appl
5	1726	100.0	332	14	US-10-084-843-53		Sequence 53, Appl
6	1726	100.0	332	14	US-10-098-732A-41		Sequence 41, Appl
7	1542	89.3	286	10	US-09-985-372-4		Sequence 4, Appli
8	1499	86.8	652	14	US-10-193-002-350		Sequence 350, App
9	1499	86.8	652	14	US-10-084-843-355		Sequence 355, App
10	1499	86.8	802	9	US-09-287-849-10		Sequence 10, Appl
11	1499	86.8	802	14	US-10-193-002-209		Sequence 209, App
12	1499	86.8	802	14	US-10-193-002-346		Sequence 346, App
13	1499	86.8	802	14	US-10-084-843-214		Sequence 214, App
14	1499	86.8	802	14	US-10-084-843-351		Sequence 351, App
15	1499	86.8	802	14	US-10-359-460-10		Sequence 10, Appl

16	248.5	14.4	273	16	US-10-437-963-184760	Sequence 184760,
17	241	14.0	402	16	US-10-437-963-141313	Sequence 141313,
18	240.5	13.9	437	16	US-10-437-963-168762	Sequence 168762,
19	239	13.8	431	16	US-10-437-963-204963	Sequence 204963,
20	238	13.8	1269	16	US-10-437-963-130985	Sequence 130985,
21	237.5	13.8	697	12	US-10-425-114-41545	Sequence 41545, A
22	233.5	13.5	547	16	US-10-437-963-102579	Sequence 102579,
23	232.5	13.5	270	16	US-10-437-963-130068	Sequence 130068,
24	232.5	13.5	360	16	US-10-437-963-179473	Sequence 179473,
25	230.5	13.4	426	16	US-10-437-963-142229	Sequence 142229,
26	230.5	13.4	584	14	US-10-156-761-12405	Sequence 12405, A
27	229.5	13.3	569	16	US-10-437-963-149276	Sequence 149276,
28	229	13.3	440	16	US-10-437-963-128104	Sequence 128104,
29	228.5	13.2	235	16	US-10-437-963-182861	Sequence 182861,
30	227.5	13.2	652	15	US-10-104-047-3364	Sequence 3364, A
31	227	13.2	598	12	US-10-282-122A-84954	Sequence 84954, A
32	226.5	13.1	504	16	US-10-437-963-128105	Sequence 128105,
33	223.5	12.9	299	16	US-10-437-963-134054	Sequence 134054,
34	223.5	12.9	501	16	US-10-437-963-161137	Sequence 161137,
35	223.5	12.9	533	16	US-10-437-963-167716	Sequence 167716,
36	221.5	12.8	556	16	US-10-437-963-137918	Sequence 137918,
37	220.5	12.8	406	16	US-10-437-963-122256	Sequence 122256,
38	220	12.7	389	12	US-10-425-114-56400	Sequence 56400, A
39	220	12.7	466	16	US-10-437-963-195119	Sequence 195119,
40	219.5	12.7	369	16	US-10-437-963-196554	Sequence 196554,
41	219.5	12.7	376	16	US-10-437-963-154970	Sequence 154970,
42	218.5	12.7	285	16	US-10-437-963-189482	Sequence 189482,
43	218.5	12.7	534	16	US-10-437-963-159995	Sequence 159995,
44	217.5	12.6	350	16	US-10-437-963-102583	Sequence 102583,
45	217.5	12.6	817	15	US-10-369-493-1813	Sequence 1813, Ap

ALIGNMENTS

RESULT 1

US-09-985-372-2
; Sequence 2, Application US/0985372
; Publication No. US20030054008A1
; GENERAL INFORMATION:
; APPLICANT: LAQUEYRERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/985,372
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/132,528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-985-372-2

Query Match 100.0%; Score 1726; DB 10; Length 325;
Best Local Similarity 100.0%; Pred. No. 8.6e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MHQVDNLTERRKGRLLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA	60
Db	1	MHQVDNLTERRKGRLLAALAIAMASASLVTVAVPATANADPEPAPVPTTAASPPSTAAA	60
Qy	61	PPAPATPVAPPPAAANTNAQGDNNAAPPADPNAPPPVIAAPNPQVRIIDNPVGGF	120
Db	61	PPAPATPVAPPPAAANTNAQGDNNAAPPADPNAPPPVIAAPNPQVRIIDNPVGGF	120
Qy	121	SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPVANDTRIVGLRDLQKLYASABA	180

```
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFGGPPVANDTRIVLGRDQKLYASAEA 180
Qy 181 TDSKAAARLGSDMGFEYMPYPTGTRINQETVSLDANGVSGSASYEYVKFSDSPKNGQIWT 240
Db 181 TDSKAAARLGSDMGFEYMPYPTGTRINQETVSLDANGVSGSASYEYVKFSDSPKNGQIWT 240
Qy 241 GVIGSPAANAPDAGPPQRFVWVWLTGANNPVDKGAALAESIRPLVAPPAPAPAPAP 300
Db 241 GVIGSPAANAPDAGPPQRFVWVWLTGANNPVDKGAALAESIRPLVAPPAPAPAPAP 300
Qy 301 APAPAPAGEVAPTPTTPTPTQRTLPA 325
Db 301 APAPAPAGEVAPTPTTPTPTQRTLPA 325

RESULT 2
US-09-985-372-3
; Sequence 3, Application US/09985372
; Publication No. US20030054008A1
; GENERAL INFORMATION:
; APPLICANT: LAQUEYERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM
; AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; TUBERCULOSIS
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/985,372
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/132,528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-985-372-3

Query Match 100.0%; Score 1726; DB 10; Length 325;
Best Local Similarity 100.0%; Pred. No. 8.6e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDPNLTRKGRLLAALAIAAMASASLVTVAVPATANADPEFAPVPTTAASPPSTAAA 60
Db 1 MHQVDPNLTRKGRLLAALAIAAMASASLVTVAVPATANADPEFAPVPTTAASPPSTAAA 60
Qy 61 PPAPATPVAPPAAANTNAGPDNAPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120
Db 61 PPAPATPVAPPAAANTNAGPDNAPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFGGPPVANDTRIVLGRDQKLYASAEA 180
Db 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFGGPPVANDTRIVLGRDQKLYASAEA 180
Qy 181 TDSKAAARLGSDMGFEYMPYPTGTRINQETVSLDANGVSGSASYEYVKFSDSPKNGQIWT 240
Db 181 TDSKAAARLGSDMGFEYMPYPTGTRINQETVSLDANGVSGSASYEYVKFSDSPKNGQIWT 240
Qy 241 GVIGSPAANAPDAGPPQRFVWVWLTGANNPVDKGAALAESIRPLVAPPAPAPAPAP 300
Db 241 GVIGSPAANAPDAGPPQRFVWVWLTGANNPVDKGAALAESIRPLVAPPAPAPAPAP 300
Qy 301 APAPAPAGEVAPTPTTPTPTQRTLPA 325
Db 301 APAPAPAGEVAPTPTTPTPTQRTLPA 325

RESULT 3
US-09-886-349A-41
; Sequence 41, Application US/09886349A
; Publication No. US20040086523A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Reed, Steven G.
; APPLICANT: Alderson, Mark
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009070US
; CURRENT APPLICATION NUMBER: US/09/886,349A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/597,796
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 60/265,737
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: DPEP
US-09-886-349A-41

Query Match 100.0%; Score 1726; DB 12; Length 332;
Best Local Similarity 100.0%; Pred. No. 8.7e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHQVDPNLTRKGRLLAALAIAAMASASLVTVAVPATANADPEFAPVPTTAASPPSTAAA 60
Db 8 MHQVDPNLTRKGRLLAALAIAAMASASLVTVAVPATANADPEFAPVPTTAASPPSTAAA 67
Qy 61 PPAPATPVAPPAAANTNAGPDNAPADPNAPPPVIAPNAPQPVRIIDNPVGGF 120
Db 68 PPAPATPVAPPAAANTNAGPDNAPADPNAPPPVIAPNAPQPVRIIDNPVGGF 127
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFGGPPVANDTRIVLGRDQKLYASAEA 180
Db 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFFGGPPVANDTRIVLGRDQKLYASAEA 187
Qy 181 TDSKAAARLGSDMGFEYMPYPTGTRINQETVSLDANGVSGSASYEYVKFSDSPKNGQIWT 240
Db 188 TDSKAAARLGSDMGFEYMPYPTGTRINQETVSLDANGVSGSASYEYVKFSDSPKNGQIWT 247
Qy 241 GVIGSPAANAPDAGPPQRFVWVWLTGANNPVDKGAALAESIRPLVAPPAPAPAPAP 300
Db 248 GVIGSPAANAPDAGPPQRFVWVWLTGANNPVDKGAALAESIRPLVAPPAPAPAPAP 307
Qy 301 APAPAPAGEVAPTPTTPTPTQRTLPA 325
Db 308 APAPAPAGEVAPTPTTPTPTQRTLPA 332

RESULT 4
US-10-193-002-53
; Sequence 53, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, David C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendricksen, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
```

CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/193,002
FILING DATE: 10-Jul-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-10-193-002-53

Query Match 100.0%; Score 1726; DB 14; Length 332;
Best Local Similarity 100.0%; Pred. No. 8.7e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHQVDNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
Db 8 MHQVDNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 67
Qy 61 PPAPATPVAPPAAAANTPNAQPGDPNAPPPVIAAPNAPQVRIIDNPVGGF 120
Db 68 PPAPATPVAPPAAAANTPNAQPGDPNAPPPVIAAPNAPQVRIIDNPVGGF 127
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGPPGVANDTRIVLGRDOKLYASAEA 180
Db 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGPPGVANDTRIVLGRDOKLYASAEA 187
Qy 181 TDSKAAARLGSMDGFEYMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 240
Db 188 TDSKAAARLGSMDGFEYMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 247
Qy 241 GVIGSPAANAPDAGPPQRFVWLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300
Db 248 GVIGSPAANAPDAGPPQRFVWLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 307
Qy 301 APAPAPAGEVAPTPTTPTPQRTPLA 325
Db 308 APAPAPAGEVAPTPTTPTPQRTPLA 332

RESULT 5
US-10-084-843-53
; Sequence 53, Application US/10084843
; Publication No. US20030143243A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
Skeiky, Yasir A.W.
Dillon, Davin C.
Campos-Neto, Antonio
Houghton, Raymond
Vedvick, Thomas S.
Twardzik, Daniel R.

Lodes, Michael J.
Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 355
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/084,843
FILING DATE: 25-Feb-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-10-084-843-53

Query Match 100.0%; Score 1726; DB 14; Length 332;
Best Local Similarity 100.0%; Pred. No. 8.7e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MHQVDNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
Db 8 MHQVDNLTTRKGRLLAALAIAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 67
Qy 61 PPAPATPVAPPAAAANTPNAQPGDPNAPPPVIAAPNAPQVRIIDNPVGGF 120
Db 68 PPAPATPVAPPAAAANTPNAQPGDPNAPPPVIAAPNAPQVRIIDNPVGGF 127
Qy 121 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGPPGVANDTRIVLGRDOKLYASAEA 180
Db 128 SFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGPPGVANDTRIVLGRDOKLYASAEA 187
Qy 181 TDSKAAARLGSMDGFEYMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 240
Db 188 TDSKAAARLGSMDGFEYMPYPGTRINQETVSLDANGVSGSASYEYVKFSDPSKPNQIWT 247
Qy 241 GVIGSPAANAPDAGPPQRFVWLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 300
Db 248 GVIGSPAANAPDAGPPQRFVWLTANNPVVDKGAALAESIRPLVAPPAPAPAPAP 307
Qy 301 APAPAPAGEVAPTPTTPTPQRTPLA 325
Db 308 APAPAPAGEVAPTPTTPTPQRTPLA 332

RESULT 6
US-10-098-732A-41
; Sequence 41, Application US/10098732A
; Publication No. US20030175294A1

```
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guiderian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 41
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:DPEP
US-10-098-732A-41

Query Match      100.0%; Score 1726; DB 14; Length 332;
Best Local Similarity 100.0%; Pred. No. 8.7e-106;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQVDPNLTTRKRLAALAAIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 60
DB 8 MEQVDPNLTTRKRLAALAAIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPSTAAA 67
QY 61 PPAPATPVAPPPAAATNPNAOPGDPNAPPPADNPAPPPVPIAPNAPQVRIQNPVGGF 120
DB 68 PPAPATPVAPPPAAATNPNAOPGDPNAPPPADNPAPPPVPIAPNAPQVRIQNPVGGF 127
QY 121 SPALPAGWVESDAHFYDGSALLSKTTGDPFPGQPPVANDTRIVLGRDQKLYASABA 180
DB 128 SPALPAGWVESDAHFYDGSALLSKTTGDPFPGQPPVANDTRIVLGRDQKLYASABA 187
QY 181 TDSKAARLGSMDGFMFYPGTRINQETVSLDANGVSGSASYEVKFSKNGQIWT 240
DB 188 TDSKAARLGSMDGFMFYPGTRINQETVSLDANGVSGSASYEVKFSKNGQIWT 247
QY 241 VIGVSPAAANAPDAGPPQRFVFWLGTANNPVDKGAALAESIRPLVAPPAPAPAPAE 300
DB 248 VIGVSPAAANAPDAGPPQRFVFWLGTANNPVDKGAALAESIRPLVAPPAPAPAPAE 307
QY 301 APAPAPAGEVAPTPTTPTPQRTLPA 325
DB 308 APAPAPAGEVAPTPTTPTPQRTLPA 332

RESULT 7
US-09-985-372-4
; Sequence 4, Application US/09985372
; Publication No. US20030054008A1
; GENERAL INFORMATION:
; APPLICANT: LAQUEYRIERIE, Anne
; APPLICANT: MARCHAL, Gilles
; APPLICANT: PESCHER, Pascale
; APPLICANT: ROMAIN, Felix
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICORORGANISMS PRODUCING THEM
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF
; FILE REFERENCE: 0660-0138-ODIV
; CURRENT APPLICATION NUMBER: US/09/985,372
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/132,528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 286
; TYPE: PRT

; ORGANISM: Mycobacterium tuberculosis
US-09-985-372-4

Query Match      89.3%; Score 1542; DB 10; Length 286;
Best Local Similarity 100.0%; Pred. No. 9.3e-94;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 DPEPAPPVPTTAASPPSTAAAAPATPVAPPPPPAAATNPNAOPGDPNAPPPADNPAPP 99
DB 1 DPEPAPPVPTTAASPPSTAAAAPATPVAPPPPPAAATNPNAOPGDPNAPPPADNPAPP 60
QY 100 PVIAENAPQVRIQNPVGGFSPALPAGWVESDAHFYDGSALLSKTTGDPFPGQPPV 159
DB 61 PVIAENAPQVRIQNPVGGFSPALPAGWVESDAHFYDGSALLSKTTGDPFPGQPPV 120
QY 160 ANDTRIVLGRDQKLYASAEATDSKAAARLGSMDGFMFYPGTRINQETVSLDANGVSG 219
DB 121 ANDTRIVLGRDQKLYASAEATDSKAAARLGSMDGFMFYPGTRINQETVSLDANGVSG 180
QY 220 SASIYEVKFSKNGQIWTGIVGSPAAANAPDAGPPQRFVFWLGTANNPVDKGAAL 279
DB 181 SASIYEVKFSKNGQIWTGIVGSPAAANAPDAGPPQRFVFWLGTANNPVDKGAAL 240
QY 280 AESIRPLVAPPAPAPAPAEPAEPAPAPAPAGEVAPTPTTPTPQRTLPA 325
DB 241 AESIRPLVAPPAPAPAEPAEPAPAPAPAGEVAPTPTTPTPQRTLPA 286

RESULT 8
US-10-193-002-350
; Sequence 350, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Devin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSES: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/193,002
; FILING DATE: 10-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 350:
; SEQUENCE CHARACTERISTICS:
```

[illegible]

```
;
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 209:
US-10-193-002-209

Query Match      86.8%; Score 1499; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 1.9e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPNAAAPPADPNAPPVPIAPN 106
DB 524 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPNAAAPPADPNAPPVPIAPN 583
QY 107 APQVRIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPVANDTRIV 166
DB 584 APQVRIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPVANDTRIV 643
QY 167 LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV 226
DB 644 LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV 703
QY 227 KFSDSPKPNQIWTGIVGSPAANAPDAGPPQRFVVMGLGTANNPVDKGAALAESIRPL 286
DB 704 KFSDSPKPNQIWTGIVGSPAANAPDAGPPQRFVVMGLGTANNPVDKGAALAESIRPL 763
QY 287 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTPLA 325
DB 764 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTPLA 802

RESULT 11
US-10-193-002-209
; Sequence 209, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
;           Skeiky, Yasir A.W.
;           Dillon, Davin C.
;           Campos-Neto, Antonia
;           Houghton, Raymond
;           Vedvick, Thomas S.
;           Twardzik, Daniel R.
;           Lodes, Michael J.
;           Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                   TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SEED and BERRY LLP
;   STREET: 6300 Columbia Center, 701 Fifth Avenue
;   CITY: Seattle
;   STATE: Washington
;   COUNTRY: USA
;   ZIP: 98104-7092
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/193,002
;   FILING DATE: 10-Jul-2002
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/09/072,596
;   FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
;   NAME: Maki, David J.
;   REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (206) 622-4900
;   TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 209:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 802 amino acids
;   TYPE: amino acid
;
;
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 209:
US-10-193-002-209

Query Match      86.8%; Score 1499; DB 14; Length 802;
Best Local Similarity 100.0%; Pred. No. 1.9e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPNAAAPPADPNAPPVPIAPN 106
DB 524 VPTTAASPPSTAAAPAPATPVAPPAAANTPNAQPGDPNAAAPPADPNAPPVPIAPN 583
QY 107 APQVRIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPVANDTRIV 166
DB 584 APQVRIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPFGQPPPVANDTRIV 643
QY 167 LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV 226
DB 644 LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMYPGTRINQETVSLDANGVSGSASYEV 703
QY 227 KFSDSPKPNQIWTGIVGSPAANAPDAGPPQRFVVMGLGTANNPVDKGAALAESIRPL 286
DB 704 KFSDSPKPNQIWTGIVGSPAANAPDAGPPQRFVVMGLGTANNPVDKGAALAESIRPL 763
QY 287 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTPLA 325
DB 764 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTPLA 802

RESULT 12
US-10-193-002-346
; Sequence 346, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
;           Skeiky, Yasir A.W.
;           Dillon, Davin C.
;           Campos-Neto, Antonia
;           Houghton, Raymond
;           Vedvick, Thomas S.
;           Twardzik, Daniel R.
;           Lodes, Michael J.
;           Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                   TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SEED and BERRY LLP
;   STREET: 6300 Columbia Center, 701 Fifth Avenue
;   CITY: Seattle
;   STATE: Washington
;   COUNTRY: USA
;   ZIP: 98104-7092
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/193,002
;   FILING DATE: 10-Jul-2002
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/09/072,596
;   FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
;   NAME: Maki, David J.
;   REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (206) 622-4900
;   TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 209:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 802 amino acids
;   TYPE: amino acid
```

```
; INFORMATION FOR SEQ ID NO: 346:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 802 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 346:
US-10-193-002-346

Query Match      86.8%; Score 1499; DB 14; Length 802;
Best Local Similarity 100.0%; Pred. No. 1.9e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQGDPAAPADPNAPPPVIAPN 106
DB 524 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQGDPAAPADPNAPPPVIAPN 583
QY 107 APOQVRIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPVANDTRIV 166
DB 584 APOQVRIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPVANDTRIV 643
QY 167 LGRLDQKLYASAEATDSKAAARLGSMDGEFYMPYPTGTRINQETVSLDANGVSGSASYEV 226
DB 644 LGRLDQKLYASAEATDSKAAARLGSMDGEFYMPYPTGTRINQETVSLDANGVSGSASYEV 703
QY 227 KFSDPKPNQGIWTVIGISPAANAPDAGPPQRFVFWLGTANNPVDKGAALAESIRPL 286
DB 704 KFSDPKPNQGIWTVIGISPAANAPDAGPPQRFVFWLGTANNPVDKGAALAESIRPL 763
QY 287 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTTPQRTPLA 325
DB 764 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTTPQRTPLA 802

RESULT 13
US-10-084-843-214
; Sequence 214, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Skeiky, Yasir A.W.
; Dillon, Davin C.
; Campos-Neto, Antonio
; Houghton, Raymond
; Vedvick, Thomas S.
; Twardzik, Daniel R.
; Lodes, Michael J.
; Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.

; INFORMATION FOR SEQ ID NO: 346:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 802 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 346:
US-10-193-002-346

Query Match      86.8%; Score 1499; DB 14; Length 802;
Best Local Similarity 100.0%; Pred. No. 1.9e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQGDPAAPADPNAPPPVIAPN 106
DB 524 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQGDPAAPADPNAPPPVIAPN 583
QY 107 APOQVRIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPVANDTRIV 166
DB 584 APOQVRIDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPFPGQPPVANDTRIV 643
QY 167 LGRLDQKLYASAEATDSKAAARLGSMDGEFYMPYPTGTRINQETVSLDANGVSGSASYEV 226
DB 644 LGRLDQKLYASAEATDSKAAARLGSMDGEFYMPYPTGTRINQETVSLDANGVSGSASYEV 703
QY 227 KFSDPKPNQGIWTVIGISPAANAPDAGPPQRFVFWLGTANNPVDKGAALAESIRPL 286
DB 704 KFSDPKPNQGIWTVIGISPAANAPDAGPPQRFVFWLGTANNPVDKGAALAESIRPL 763
QY 287 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTTPQRTPLA 325
DB 764 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTTPQRTPLA 802

RESULT 14
US-10-084-843-351
; Sequence 351, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Skeiky, Yasir A.W.
; Dillon, Davin C.
; Campos-Neto, Antonio
; Houghton, Raymond
; Vedvick, Thomas S.
; Twardzik, Daniel R.
; Lodes, Michael J.
; Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
```

```
/ APPLICATION NUMBER: US/09/072,967
/ FILING DATE: 05-MAY-1998
/ ATTORNEY/AGENT INFORMATION:
/   NAME: Maki, David J.
/   REGISTRATION NUMBER: 31,392
/   REFERENCE/DOCKET NUMBER: 210121.411C9
/ TELECOMMUNICATION INFORMATION:
/   TELEPHONE: (206) 622-4900
/   TELEFAX: (206) 682-6031
/ INFORMATION FOR SEQ ID NO: 351:
/   SEQUENCE CHARACTERISTICS:
/     LENGTH: 802 amino acids
/     TYPE: amino acid
/     STRANDEDNESS: single
/     TOPOLOGY: linear
/     MOLECULE TYPE: protein
/   SEQUENCE DESCRIPTION: SEQ ID NO: 351:
US-10-084-843-351

Query Match      86.8%; Score 1499; DB 14; Length 802;
Best Local Similarity 100.0%; Pred. No. 1.9e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAAADPNAPPPVIAPN 106
Db 524 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAAADPNAPPPVIAPN 583
QY 107 APOQVRIDNPVGGFSFALPAGWVESDAAHFYDGSALLSKTTGDPFPFGQPPPVANDTRIV 166
Db 584 APOQVRIDNPVGGFSFALPAGWVESDAAHFYDGSALLSKTTGDPFPFGQPPPVANDTRIV 643
QY 167 LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMPYPTGTRINQETVSLDANGVSGSASYEV 226
Db 644 LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMPYPTGTRINQETVSLDANGVSGSASYEV 703
QY 227 KFSDPKPNQGIWTGVIIGSPAANADAGPPQRFVFWLGTANNPVDKGAAKALAESIRPL 286
Db 704 KFSDPKPNQGIWTGVIIGSPAANADAGPPQRFVFWLGTANNPVDKGAAKALAESIRPL 325
QY 287 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTLPA 325
Db 764 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTLPA 802

Search completed: July 7, 2004, 18:33:52
Job time : 61.5 secs

/ APPLICATION NUMBER: US/09/072,967
/ FILING DATE: 05-MAY-1998
/ ATTORNEY/AGENT INFORMATION:
/   NAME: Maki, David J.
/   REGISTRATION NUMBER: 31,392
/   REFERENCE/DOCKET NUMBER: 210121.411C9
/ TELECOMMUNICATION INFORMATION:
/   TELEPHONE: (206) 622-4900
/   TELEFAX: (206) 682-6031
/ INFORMATION FOR SEQ ID NO: 351:
/   SEQUENCE CHARACTERISTICS:
/     LENGTH: 802 amino acids
/     TYPE: amino acid
/     STRANDEDNESS: single
/     TOPOLOGY: linear
/     MOLECULE TYPE: protein
/   SEQUENCE DESCRIPTION: SEQ ID NO: 351:
US-10-084-843-351

Query Match      86.8%; Score 1499; DB 14; Length 802;
Best Local Similarity 100.0%; Pred. No. 1.9e-90;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAAADPNAPPPVIAPN 106
Db 524 VPTTAASPPSTAAAPAPATPVAPPPAAANTPNAQPGDPNAAADPNAPPPVIAPN 583
QY 107 APOQVRIDNPVGGFSFALPAGWVESDAAHFYDGSALLSKTTGDPFPFGQPPPVANDTRIV 166
Db 584 APOQVRIDNPVGGFSFALPAGWVESDAAHFYDGSALLSKTTGDPFPFGQPPPVANDTRIV 643
QY 167 LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMPYPTGTRINQETVSLDANGVSGSASYEV 226
Db 644 LGRLDQKLYASAEATDSKAAARLGSDMGFEFYMPYPTGTRINQETVSLDANGVSGSASYEV 703
QY 227 KFSDPKPNQGIWTGVIIGSPAANADAGPPQRFVFWLGTANNPVDKGAAKALAESIRPL 286
Db 704 KFSDPKPNQGIWTGVIIGSPAANADAGPPQRFVFWLGTANNPVDKGAAKALAESIRPL 325
QY 287 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTLPA 325
Db 764 VAPPPAPAPAPAPAPAPAPAGEVAPTPTTPTPQRTLPA 802

RESULT 15
US-10-359-460-10
/ Sequence 10, Application US/10359460
/ Publication NO. US20030147911A1
/ GENERAL INFORMATION:
/   APPLICANT: Reed, Steven G.
/   APPLICANT: Skeiky, Yasir A.W.
/   APPLICANT: Dillon, Davin C.
/   APPLICANT: Alderson, Mark
/   APPLICANT: Campos-Neto, Antonio
/   APPLICANT: Corixa Corporation
/   TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
/   TITLE OF INVENTION: and Their Uses
/   FILE REFERENCE: 014058-009020US
/   CURRENT APPLICATION NUMBER: US/10/359,460
/   CURRENT FILING DATE: 2003-02-05
/   PRIOR APPLICATION NUMBER: US/09/287,849
/   PRIOR FILING DATE: 1999-04-07
/   PRIOR APPLICATION NUMBER: US 08/818,112
/   PRIOR FILING DATE: 1997-03-13
/   PRIOR APPLICATION NUMBER: US 08/942,578
/   PRIOR FILING DATE: 1997-10-01
/   PRIOR APPLICATION NUMBER: US 09/025,197
/   PRIOR FILING DATE: 1998-02-18
/   PRIOR APPLICATION NUMBER: US 09/056,556
/   PRIOR FILING DATE: 1998-04-07
/   PRIOR APPLICATION NUMBER: US 09/223,040
/   PRIOR FILING DATE: 1998-12-30
/   NUMBER OF SEQ ID NOS: 46
```